

STIC-Biotech/ChemLib

112464

From: Whiteman, Brian
Sent: Wednesday, January 21, 2004 8:05 AM
To: STIC-Biotech/ChemLib
Subject: Fw: seq search

Labelled mailbox and room number

09/807,802
Wilson et al., 11/29/01

search SEQ ID NOs: 13, 15, and 17 against us patent and us patent application databases.

the sequences are AAV-1 capsid proteins

Thanks,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

-----Original Message-----

From: Whiteman, Brian
Sent: Wednesday, January 21, 2004 7:36 AM
To: STIC-Biotech/ChemLib
Subject: updated seq search

09/807,802
Wilson et al., 11/29/01

search SEQ ID NOs: 13, 15, and 17 against us patent and us patent application databases.

the sequences are AAV-1 capsid proteins

Thank you,

Brian Whiteman, 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
Remsen, 2D14
(703) 305-0775

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

RECEIVED
JAN 21 2004
STIC

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:00:06 ; Search time 38.198 Seconds
(without alignments)
4972.168 Million cell updates/sec

Title: US-09-807-802A-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLENDLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3989	100.0	736	12	Q9WBP8 adeno-assoc
2	3963	99.3	736	12	O56137 adeno-assoc
3	3511	88.0	736	12	O56139 adeno-assoc
4	3494	87.6	736	12	O65311 adeno-assoc
5	3442.5	86.3	737	12	Q8JQG0 adeno-assoc
6	3405	85.4	738	12	Q8JQF8 adeno-assoc
7	3402.5	85.3	735	12	O56652 adeno-assoc
8	2759.5	69.2	598	12	O56653 adeno-assoc
9	2486.5	62.3	734	12	O41855 adeno-assoc
10	2481.5	62.2	533	12	O92917 adeno-assoc
11	2220	55.7	724	12	Q9YIJ1 adeno-assoc
12	2190.5	54.9	732	12	Q83290 muscovy duc
13	2173.5	54.5	732	12	Q67666 goose parvo
14	2168.5	54.4	732	12	Q83289 muscovy duc
15	2165.5	54.3	732	12	Q8V395 goose parvo
16	2162.5	54.2	732	12	Q65444 barbarie du

17	1990.5	49.9	676	12	Q67672	Q67672	goose parvo
18	1712	42.9	587	12	Q67667	Q67667	goose parvo
19	1698	42.6	534	12	Q67668	Q67668	goose parvo
20	1695	42.5	587	12	Q65445	Q65445	barbarie du
21	1676	42.0	534	12	Q65446	Q65446	barbarie du
22	859.5	21.5	703	12	Q8QQV5	Q8QQV5	minute viru
23	715	17.9	179	12	Q9W8U7	Q9W8U7	duck parvov
24	715	17.9	179	12	Q9WA24	Q9WA24	goose parvo
25	711	17.8	179	12	Q9WN18	Q9WN18	duck parvov
26	709	17.8	179	12	Q9WN19	Q9WN19	duck parvov
27	707	17.7	179	12	Q9WN20	Q9WN20	goose parvo
28	618	15.5	571	12	Q8QQV4	Q8QQV4	minute viru
29	617	15.5	785	12	Q9J0X4	Q9J0X4	pig-tailed
30	614.5	15.4	781	12	Q9PZT0	Q9PZT0	human parvo
31	614	15.4	781	12	Q8JN56	Q8JN56	human parvo
32	611	15.3	981	12	Q8QVL1	Q8QVL1	hamster par
33	608.5	15.3	781	12	Q9JGP8	Q9JGP8	human parvo
34	606.5	15.2	781	12	P89318	P89318	human parvo
35	606.5	15.2	781	12	Q912B8	Q912B8	human eryth
36	606.5	15.2	781	12	P89319	P89319	human parvo
37	606.5	15.2	781	12	P89317	P89317	human parvo
38	605.5	15.2	781	12	Q8JYE3	Q8JYE3	erythroviru
39	605	15.2	781	12	Q8JYD9	Q8JYD9	erythroviru
40	604.5	15.2	773	12	Q913X1	Q913X1	human parvo
41	604.5	15.2	781	12	P89316	P89316	human parvo
42	603.5	15.1	781	12	P90223	P90223	human parvo
43	603.5	15.1	781	12	P90221	P90221	human parvo
44	603.5	15.1	781	12	P90224	P90224	human parvo
45	603.5	15.1	781	12	Q85117	Q85117	human parvo

ALIGNMENTS

RESULT 1

Q9WBP8 PRELIMINARY; PRT; 736 AA.
AC Q9WBP8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=85106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99214338; PubMed=10196295;
RA Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;
RT "Gene therapy vectors based on adeno-associated virus type 1.";
RL J. Virol. 73:3994-4003(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Xiao W., Wilson J.M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF063497; AAD27757.1; -.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat; 1.
SQ SEQUENCE 736 AA; 81375 MW; CFABFB9BD5CD0595 CRC64;

Query Match 100.0%; Score 3989; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 2e-266;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAADGYLPDWLENDLSEGRWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPENGLD	60
Db	1	MAADGYLPDWLENDLSEGRWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPENGLD	60
QY	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERLQEDTSFGGNLGRAVFO	120
Db	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERLQEDTSFGGNLGRAVFO	120
QY	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGTQQQPAKKRLNFGQTGDSE	180

db	121	AKRVLPLGLVEEAKTAPGKKRPVEQSPQEPDSSSGIGKTQQPAKKRLNFGQTGDSE	180
QY	181	SVPDPQPLGEPATPAAVGPTTMASSGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240
Db	181	SVPDPQPLGEPATPAAVGPTTMASSGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240
QY	241	TTSTRTWALPTYNHLYKQISSASTGASNDNHFGYSTPWGYFDENRFCHFSPRDWQRL	300
Db	241	TTSTRTWALPTYNHLYKQISSASTGASNDNHFGYSTPWGYFDENRFCHFSPRDWQRL	300
QY	301	INNNWGFRPKLNFKLFNIQVKEVTTNDGVTTIANNLTSVQVFSDSYQLPYVLGSAHQ	360
Db	301	INNNWGFRPKLNFKLFNIQVKEVTTNDGVTTIANNLTSVQVFSDSYQLPYVLGSAHQ	360
QY	361	GCLPPFPADVFMIPOGYLTLNNGSQAVGRSSFYCYLEFPSPQMLRTGNFTFSYTFEEVP	420
Db	361	GCLPPFPADVFMIPOGYLTLNNGSQAVGRSSFYCYLEFPSPQMLRTGNFTFSYTFEEVP	420
QY	421	FHSSYAHQSOLDRLMNLPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLP	480
Db	421	FHSSYAHQSOLDRLMNLPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLP	480
QY	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGV	540
Db	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGV	540
QY	541	MIFGESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG	600
Db	541	MIFGESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG	600
QY	601	ALPGMWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKNPPPOILLIKNTVPANPPA	660
Db	601	ALPGMWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKNPPPOILLIKNTVPANPPA	660
QY	661	EFSA TKFASFIQTSTGQVSVEIEWELQKENS KRWNPVEQYTSNYAKSANVDFTVDNNGL	720
Db	661	EFSA TKFASFIQTSTGQVSVEIEWELQKENS KRWNPVEQYTSNYAKSANVDFTVDNNGL	720
QY	721	YTEPRPIGTRYLTRPL	736
Db	721	YTEPRPIGTRYLTRPL	736

```

RESULT 2
O56137      PRELIMINARY;      PRT;      736 AA.
ID  O56137
AC  O56137;
DT  01-JUN-1998 (TrEMBLrel. 06, Created)
DT  01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT  01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE  Capsid protein VP1.
OS  Adeno-associated virus 6.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX  NCBI_TaxID=68558;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98080418; PubMed=9420229;
RA  Rutledge E.A.; Halbert C.L.; Russell D.W.;
RT  "Infectious clones and vectors derived from adeno-associated virus
RL  (AAV) serotypes other than AAV type 2.";
RN  J. Virol. 72:309-319(1998).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Rutledge E.A.; Russell D.W.;
RL  Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF028704; AAB95450.1; -.
DR  InterPro; IPR001403; Parvo_coat.
DR  Pfam; PF00740; Parvo_coat; 1.
SQ  SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match      99.3%; Score 3963; DB 12; Length 736;
Rest Local Similarity 99.2%; Pred. No. 1.2e-264;

```

Matches	730;	Conservative	3;	Mismatches	3;	Indels	0;	Gaps	0;		
QY	1	MAADGYLPDWLE	DLNLSE	GIREW	DLKPGAPK	ANQQKQ	DDGRGLVLP	PGYKYL	GPFG	60	
Db	1	MAADGYLPDWLE	DLNLSE	GIREW	DLKPGAPK	ANQQKQ	DDGRGLVLP	PGYKYL	GPFG	60	
QY	61	KGEPVNAADAA	ALEHDKAYDQ	QQLKAGD	NPYLRYNH	ADAEFO	ERLQED	TSFG	GNLGRA	120	
Db	61	KGEPVNAADAA	ALEHDKAYDQ	QQLKAGD	NPYLRYNH	ADAEFO	ERLQED	TSFG	GNLGRA	120	
QY	121	AKKRVLEPLGL	VEEGAKTAPG	KKRPVEQ	SPQEPD	SSGIGKTG	QQQPAK	KRLNFG	QTGDSE	180	
Db	121	AKKRVLEPLGL	VEEGAKTAPG	KKRPVEQ	SPQEPD	SSGIGKTG	QQQPAK	KRLNFG	QTGDSE	180	
QY	181	SVPDPQPLGEP	PATPAAVG	PTTMA	SGGAPMA	DNNNEG	DVG	NASGN	WHCDST	240	
Db	181	SVPDPQPLGEP	PATPAAVG	PTTMA	SGGAPMA	DNNNEG	DVG	NASGN	WHCDST	240	
QY	241	TTSTRTWALPT	VNNHLYKQ	ISSAST	GSNDNH	FGYSTP	WGYFD	FNRFCH	FSFPRD	300	
Db	241	TTSTRTWALPT	VNNHLYKQ	ISSAST	GSNDNH	FGYSTP	WGYFD	FNRFCH	FSFPRD	300	
QY	301	INNNWGFRPK	RNFKLFNI	QVKEVT	TNDGVTTI	ANNLTST	VQVFS	DSEYQL	PVVLGSA	360	
Db	301	INNNWGFRPK	RNFKLFNI	QVKEVT	TNDGVTTI	ANNLTST	VQVFS	DSEYQL	PVVLGSA	360	
QY	361	GCLPPFPADV	FMIPOGYLT	LNNGSQA	VGRSSF	CYLEYFP	SQMLRT	GNNFT	FSYTFEE	420	
Db	361	GCLPPFPADV	FMIPOGYLT	LNNGSQA	VGRSSF	CYLEYFP	SQMLRT	GNNFT	FSYTFEE	420	
QY	421	FHSSYAH	SQSLDRLM	NPLIDQ	LYLNR	TQNSG	SAQNK	DLLFSR	GSPAG	MSVP	480
Db	421	FHSSYAH	SQSLDRLM	NPLIDQ	LYLNR	TQNSG	SAQNK	DLLFSR	GSPAG	MSVP	480
QY	481	GPCYRQ	RVS	KTKTDNN	SNFTWT	GASKYN	LNGRES	IINPG	TAMASH	KDDK	540
Db	481	GPCYRQ	RVS	KTKTDNN	SNFTWT	GASKYN	LNGRES	IINPG	TAMASH	KDDK	540
QY	541	MIFGES	SAGAS	TALD	NVMITD	EEEEKAT	NPVATER	FGTVAV	NFQSS	SDPATG	600
Db	541	MIFGES	SAGAS	TALD	NVMITD	EEEEKAT	NPVATER	FGTVAV	NFQSS	SDPATG	600
QY	601	ALPGMW	QDRDVY	LGQPI	WAKIPH	TDGH	FHPSP	LMGG	FLKNPP	QIILIK	660
Db	601	ALPGMW	QDRDVY	LGQPI	WAKIPH	TDGH	FHPSP	LMGG	FLKNPP	QIILIK	660
QY	661	EFSATK	FASFI	TOYSTG	QVSVEI	EWE	LQENSK	RWNPE	VQYTSN	YAKSAN	720
Db	661	EFSATK	FASFI	TOYSTG	QVSVEI	EWE	LQENSK	RWNPE	VQYTSN	YAKSAN	720
QY	721	YTEPRPIG	TRYL	TRPL	736						
Db	721	YTEPRPIG	TRYL	TRPL	736						

```

RESULT 3
O56139 PRELIMINARY; PRT; 736 AA.
ID O56139
AC O56139;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; Pubmed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).

```

RN	[2]	DB	Capsid protein.	OS	Adeno-associated virus 3.
RP	SEQUENCE FROM N.A.	OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	OX	NCBI_TaxID=46350;
RA	Rutledge E.A., Russell D.W.;	RN	[1]	RP	SEQUENCE FROM N.A.
RL	Submitted (SSP-1997) to the EMBL/GenBank/DBJ databases.	RC	STRAIN=3H;	RX	MEDLINE=96266430; PubMed=8661429;
DR	EMBL; AF028705; AAB95452.1; -.	RA	Muramatsu S., Brown K.E.;	RA	Muramatsu S., Mizukami H., Young N.S., Brown K.E.;
DR	InterPro; IPR001403; Parvo_coat.	RT	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.	RT	"Nucleotide sequencing and generation of an infectious clone of adeno-
DR	Pfam; PF00740; Parvo_coat; 1.	RT	associated virus 3."	RL	Virology 221:208-217(1996).
SQ	SEQUENCE 736 AA; 81906 MW; DD52331AD5F0D70F CRC64;	RN	[2]	RP	SEQUENCE FROM N.A.
		RC	STRAIN=3H;	RC	STRAIN=3H;
		RA	Muramatsu S., Brown K.E.;	RA	Muramatsu S., Brown K.E.;
		RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.	RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
		DR	EMBL; U48704; AAC55049.1; -.	DR	EMBL; U48704; AAC55049.1; -.
		DR	InterPro; IPR001403; Parvo_coat.	DR	InterPro; IPR001403; Parvo_coat.
		DR	Pfam; PF00740; Parvo_coat; 1.	DR	Pfam; PF00740; Parvo_coat; 1.
		SQ	SEQUENCE 736 AA; 81660 MW; AFF1EF47B5C67A10 CRC64;	SQ	SEQUENCE 736 AA; 81660 MW; AFF1EF47B5C67A10 CRC64;
			Query Match 87.6%; Score 3494; DB 12; Length 736;		Query Match 87.6%; Score 3494; DB 12; Length 736;
			Best Local Similarity 86.4%; Pred. No. 2.6e-232;		Best Local Similarity 86.4%; Pred. No. 2.6e-232;
			Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;		Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;
Qy	1	Qy	1	Qy	1
Db	1	Db	1	Db	1
Qy	61	Qy	61	Qy	61
Db	61	Db	61	Db	61
Qy	121	Qy	121	Qy	121
Db	121	Db	121	Db	121
Qy	181	Qy	181	Qy	181
Db	181	Db	181	Db	181
Qy	241	Qy	241	Qy	241
Db	241	Db	241	Db	241
Qy	301	Qy	301	Qy	301
Db	300	Db	300	Db	300
Qy	361	Qy	361	Qy	361
Db	360	Db	360	Db	360
Qy	421	Qy	421	Qy	421
Db	420	Db	420	Db	420
Qy	480	Qy	480	Qy	480
Db	480	Db	480	Db	480
Qy	540	Qy	540	Qy	540
Db	540	Db	540	Db	540
Qy	600	Qy	600	Qy	600
Db	600	Db	600	Db	600
Qy	660	Qy	660	Qy	660
Db	660	Db	660	Db	660

RESULT 4
Q65311
ID Q65311
AC Q65311;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)


```
QY 720 LYTEPRPIGTRYLTRPL 736
Db :|||||
720 VYSEPRPIGTRYLTRNL 736

RESULT 5
Q8JQGO PRELIMINARY; PRT; 737 AA.
AC Q8JQGO;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 7.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=202812;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12192090;
RA Gao G.P., Alvira M.R., Wang L., Calcedo R., Johnston J., Wilson J.M.;
RT "Novel adeno-associated viruses from rhesus monkeys as vectors for
human gene therapy.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:11854-11859(2002).
DR EMBL; AF513851; AAN03855.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 737 AA; 81652 MW; D63B8CE75831140D0 CRC64;

Query Match 86.3%; Score 3442.5; DB 12; Length 737;
Best Local Similarity 85.3%; Pred. No. 9.4e-229;
Matches 630; Conservative 50; Mismatches 54; Indels 5; Gaps 4;

QY 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db :|||||
1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120
Db :|||||
61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKRPEVQSPQ-EPDSSSGIGKTGQPAKKRLNFGQTGDS 179
Db :|||||
121 AKKRVLEPLGLVEEGAKTAPGKRPEVQSPQSPDSSTGIGKKGQPAKKRLNFGQTGDS 180

QY 180 ESVPDPQPLGEPATPAAVPTTASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRV 239
Db :|||||
181 ESVPDPQPLGEPAPSSVSGTVAAGGAPMADNNEGADGVGNASGNWHCDSTWLGDRV 240

QY 240 ITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDNFRFCHFSRPDWQR 299
Db :|||||
241 ITTSTRTWALPTYNNHLYKQISSSETAGSTNDNTYFGYSTPWGYFDNFRFCHFSRPDWQR 300

QY 300 LINNNWGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLSTVQVFSDSYQLPYVLGSAH 359
Db :|||||
301 LINNNWGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLSTVQVFSDSYQLPYVLGSAH 360

QY 360 QGCLPPFPADVFMIPOGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEV 419
Db :|||||
361 QGCLPPFPADVFMIPOGYLTLNNGSQSVGRSSFYCLEYFPPSQMLRTGNNFESYFSDV 420

QY 420 PFHSSYAHQSGLDRLMNPLIDQLYLYLNRTQ-NQSGSAQNKDLFSRGSAGMSVQPKNW 478
Db :|||||
421 PFHSSYAHQSGLDRLMNPLIDQLYLYLARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNW 480

QY 479 LPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMS 538
Db :|||||
481 LPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKODEDFPSS 540

QY 539 GVMTPKESAGASN-TALDNTVMTDEEIKATNPVATERFTVAVNFQSSSTDPAQDVH 597
Db :|||||
541 GVLIIFGK--TGATNKTTLLENVLMTNEEIRPTNPVATEEYGIIVSSNLQAANTAAQTQVNV 598
```

```
QY 598 AMGALPGMVWQDRDYLQGPWIAKIPHTDGHFHPSPMLGGFGLKNNPPQILIKNTVPAN 657
Db :|||||
599 NQGALPGMVWQNRDYLQGPWIAKIPHTDGNFHPSPMLGGFGLKHPPPQILIKNTVPAN 658

QY 658 PPAEFSATKPFASFITQYSTQGVSVIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDN 717
Db :|||||
659 PPEVFTPAKPFASFITQYSTQGVSVIEWELQKENSKRWNPEIQYTSNPEKQTVGVDFAVDS 718

QY 718 NGLYTEPRPIGTRYLTRPL 736
Db :|||||
719 QGVYSEPRPIGTRYLTRNL 737

RESULT 6
Q8JQF8 PRELIMINARY; PRT; 738 AA.
AC Q8JQF8;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 8.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=202813;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12192090;
RA Gao G.P., Alvira M.R., Wang L., Calcedo R., Johnston J., Wilson J.M.;
RT "Novel adeno-associated viruses from rhesus monkeys as vectors for
human gene therapy.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:11854-11859(2002).
DR EMBL; AF513852; AAN03857.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 738 AA; 81756 MW; 59BC2BE1B148AA14 CRC64;

Query Match 85.4%; Score 3405; DB 12; Length 738;
Best Local Similarity 84.0%; Pred. No. 3.6e-226;
Matches 620; Conservative 48; Mismatches 68; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db :|||||
1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120
Db :|||||
61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKRPEVQSPQ-EPDSSSGIGKTGQPAKKRLNFGQTGDS 179
Db :|||||
121 AKKRVLEPLGLVEEGAKTAPGKRPEVQSPQSPDSSTGIGKKGQPAKKRLNFGQTGDS 180

QY 180 ESVPDPQPLGEPATPAAVPTTASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRV 239
Db :|||||
181 ESVPDPQPLGEPAPSGVGPNTMAAGGAPMADNNEGADGVSSSGNWHCDSTWLGDRV 240

QY 240 ITTSTRTWALPTYNNHLYKQISS-ASTGASNDNHYFGYSTPWGYFDNFRFCHFSRPDWQ 298
Db :|||||
241 ITTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDNFRFCHFSRPDWQ 300

QY 299 RLINNNWGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLSTVQVFSDSYQLPYVLGSA 358
Db :|||||
301 RLINNNWGFRPKRLNFKLFNIOQKEVTQNEGKTIANNLSTIQVFTDSEYQLPYVLGSA 360

QY 359 HQGCLPPFPADVFMIPOGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEE 418
Db :|||||
361 HQGCLPPFPADVFMIPOGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTYTFED 420

QY 419 VPFHSSYAHQSGLDRLMNPLIDQLYLYLNRTQNSGSAQNKOLLFSRGSAGMSVQPKNW 478
Db :|||||
421 VPFHSSYAHQSGLDRLMNPLIDQLYLYLSRTQTGGTANTQTGLFSGGPNNTMANQAKNW 480

QY 479 LPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMS 538
```

Db	481	LPGPCYRQVRSTTTGQNNNSFAWTAGTKYHLNGRNSLANPGIAMATHKDDERFFPSN	540	Db	181	SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVI	240
Qy	539	GVMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA	598	Qy	241	TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRL	300
Db	541	GILIFGKQNAARDNADYSDVMLTSEEEIKTNPVATEEYGIVADNLQQONTAPQIGTVNS	600	Db	241	TTSTRTWALPTYNNHLYKQISSQS-GASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRL	299
Qy	599	MGALPGMWQDRDVLQGPWAKIPHTDGHFHPSPMLMGGFGLKNPPPIQLIKNTVPANP	658	Qy	301	INNNGFRPKRLNFKLFNIQVKEVTNDGVTTIANNLTSVQVFSSEYQLPYVLGSAHQ	360
Db	601	QGALPGMWQNRDVLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPIQLIKNTVPADP	660	Db	300	INNNGFRPKRLNFKLFNIQVKEVTNDGVTTIANNLTSVQVFTDSEYQLPYVLGSAHQ	359
Qy	659	PAEFSATKPFASFITQYSTGVQSVSEIEWELQKNSKRWNPVQYTSNYAKSANVDTVDNN	718	Qy	361	GCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSPYCLEYFPPSQMLRTGNNFTFSYTFEEVP	420
Db	661	PTTFNQSKLNSFITQYSTGVQSVSEIEWELQKNSKRWNPVQYTSNYKSTSVDFAVNTE	720	Db	360	GCLPPFPADVFMVPQYGYLTNLNGSQAVGRSSPYCLEYFPPSQMLRTGNNFTFSYTFEDVP	419
Qy	719	GLYTEPRPIGTRYLTRPL	736	Qy	421	FHSSYAHSQSLDRLMNPPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLP	480
Db	721	GVYSEPRPIGTRYLTRNL	738	Db	420	FHSSYAHSQSLDRLMNPPLIDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLP	479
RESULT 7				RESULT 8			
O56652	PRELIMINARY; PRT; 735 AA.			O56653	PRELIMINARY; PRT; 598 AA.		
ID	O56652			ID	O56653		
AC	O56652;			AC	O56653;		
DT	01-JUN-1998 (TrEMBLrel. 06, Created)			DT	01-JUN-1998 (TrEMBLrel. 06, Created)		
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	Major coat protein VP1.			DE	Major coat protein VP2.		
OS	Adeno-associated virus 2 (AAV2).			OS	Adeno-associated virus 2 (AAV2).		
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.			OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
OX	NCBI_TaxID=10804;			OX	NCBI_TaxID=10804;		
RN	[1]			RN	[1]		
RP	SEQUENCE FROM N.A.			RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95088582; PubMed=7996133;			RX	MEDLINE=95088582; PubMed=7996133;		
RA	Ruffing M., Heid H., Kleinschmidt J.A.;			RA	Ruffing M., Heid H., Kleinschmidt J.A.;		
RT	"Mutations in the carboxy terminus of adeno-associated virus 2 capsid			RT	"Mutations in the carboxy terminus of adeno-associated virus 2 capsid		
RT	proteins affect viral infectivity: lack of an RGD integrin-binding			RT	proteins affect viral infectivity: lack of an RGD integrin-binding		
RT	motif.";			RT	motif.";		
RL	J. Gen. Virol. 75:0-0(0).			RL	J. Gen. Virol. 75:0-0(0).		
RN	[2]			RN	[2]		
RP	SEQUENCE FROM N.A.			RP	SEQUENCE FROM N.A.		
RA	Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,			RA	Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,		
RA	Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,			RA	Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,		
RA	Tratschin J.-D., Weitz M.;			RA	Tratschin J.-D., Weitz M.;		
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.			RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF043303; AAC03780.1; -.			DR	EMBL; AF043303; AAC03778.1; -.		
DR	InterPro; IPR001403; Parvo_coat.			DR	InterPro; IPR001403; Parvo_coat.		
DR	Pfam; PF00740; Parvo_coat; 1.			DR	Pfam; PF00740; Parvo_coat; 1.		
KW	Coat protein.			KW	Coat protein.		
FT	VARIANT 76 76 D -> V.			FT	VARIANT 416 416 D -> N.		
FT	VARIANT 553 553 D -> N.			FT	VARIANT 430 430 T -> N.		
FT	VARIANT 567 567 T -> N.			FT	VARIANT 540 541 QV -> HV.		
FT	VARIANT 677 678 QV -> HV.			FT	VARIANT 573 573 V -> R.		
FT	VARIANT 710 710 V -> R.						
SQ	SEQUENCE 735 AA; 81944 MW; 980BEEF46908390B CRC64;						
Query Match				Query Match			
Best Local Similarity 85.3%; Score 3402.5; DB 12; Length 735;				Best Local Similarity 83.3%; Score 3402.5; DB 12; Length 735;			
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;				Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;			
Qy	1	MAADGYLPDWNLDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD	60	Qy	1	MAADGYLPDWNLDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWNLDNLSEGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60	Db	1	MAADGYLPDWNLDNLSEGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Qy	61	KGEPVNAADAAAEHDKAYDQQLKAGDNPLYRYNHADAEEFQERLQEDTSFGGNLGRAVFQ	120	Qy	61	KGEPVNAADAAAEHDKAYDQQLKAGDNPLYRYNHADAEEFQERLQEDTSFGGNLGRAVFQ	120
Db	61	KGEPVNEADAAAEHDKAYDRQLDSDGNPLYLKNHADAEEFQERLKEDTSFGGNLGRAVFQ	120	Db	61	KGEPVNEADAAAEHDKAYDRQLDSDGNPLYLKNHADAEEFQERLKEDTSFGGNLGRAVFQ	120
Qy	121	AKKRVLEPLGLVEEAGKAPGKKRPVEQSPQEPDSSSGIGTKTQGPQAKKRLNFGQTGDSE	180	Qy	121	AKKRVLEPLGLVEEAGKAPGKKRPVEQSPQEPDSSSGIGTKTQGPQAKKRLNFGQTGDSE	180
Db	121	AKKRVLEPLGLVEEAPVKAPGKKRPVEHSPVEPDSSSGTGKAGQGPARKRLNFGQTGDAD	180	Db	121	AKKRVLEPLGLVEEAPVKAPGKKRPVEHSPVEPDSSSGTGKAGQGPARKRLNFGQTGDAD	180
Qy	181	SVPDPQPLGEPATPAAGVPTTMSAGGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240	Qy	181	SVPDPQPLGEPATPAAGVPTTMSAGGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240

SQ	SEQUENCE	598 AA; 66619 MW; 0708111ED9368E934 CRC64;
Query Match 69.2%; Score 2759.5; DB 12; Length 598;		
Best Local Similarity 82.8%; Pred. No. 8.3e-182;		
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;		
QY	139	APGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGTGDSSEVPDPQLGEPATPAAV 198
Db	2	APGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGTGDAVSVPDPQLGQPPAASGL 61
QY	199	GPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYK 258
Db	62	GNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNHLYK 121
QY	259	QISSASTGASNDNHFGYSTPWGYFDENRFHCHFSRPDWRQLINNNWGFRPKRLNFKLEN 318
Db	122	QISSQS-GASNDNHFGYSTPWGYFDENRFHCHFSRPDWRQLINNNWGFRPKRLNFKLEN 180
QY	319	IQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 378
Db	181	IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 240
QY	379	LTLNNGSQAVGRSSFYCLEYFPFSQMLRTGNFTFSYTEEVPFHSSYAHSSQSLDRLMNPL 438
Db	241	LTLNNGSQAVGRSSFYCLEYFPFSQMLRTGNFTFSYTEEVPFHSSYAHSSQSLDRLMNPL 300
QY	439	IDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNN 498
Db	301	IDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNN 360
QY	499	SNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGVMIFGKESAGASNTALDNV 558
Db	361	SEYSWTGATKYHLNGRDSLVPNGPAMASHKDEEKFFPQSGVLIFGKQGSEKTNVDIEKV 420
QY	559	MITDEEIKATNPVATERFGTVAVNPFQSSSTDPATGDVHAMGALPGMWQDRDVYLGPI 618
Db	421	MITDEEIRTNPVATEQYQGSVSTNLQRNQAATADVNTQGVLPGMVWQDRDVYLGPI 480
QY	619	WAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTGQ 678
Db	481	WAKIPHTDGHFHPSPMLGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQ 540
QY	679	VSVEIEWELQENSKRWNPVQYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL 736
Db	541	VSVEIEWELQENSKRWNPQIQTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 598
RESULT 9		
O41855	PRELIMINARY; PRT; 734 AA.	
ID	O41855	
AC	O41855;	
DT	01-JAN-1998 (TrEMBLrel. 05, Created)	
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE	Capsid.	
OS	Adeno-associated virus 4.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	
OX	NCBI_TaxID=57579;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC VR-646;	
RX	MEDLINE=97404695; PubMed=9261407;	
RA	Chiorini J.A., Yang L., Liu Y., Safer B., Kotin R.M.;	
RT	"Cloning of adeno-associated virus type 4 (AAV4) and generation of	
RT	recombinant AAV4 particles."	
RL	J. Virol. 71:6823-6833 (1997).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC VR-646;	
RA	Chiorini J.A., Yang L., Kotin R.M., Safer B.;	
RL	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; U89790; AAC58045.1; -.	
DR	InterPro; IPR001403; Parvo_coat.	

DR	Pfam; PF00740; Parvo_coat; 1.
SQ	SEQUENCE 734 AA; 80639 MW; 616CC27A777BBE6F CRC64;
Query Match 62.3%; Score 2486.5; DB 12; Length 734;	
Best Local Similarity 63.5%; Pred. No. 7.4e-163;	
Matches 475; Conservative 81; Mismatches 161; Indels 31; Gaps 10;	
QY	4 DGYLPDWLEDNLSGIREWDLKPAPKPKANQQKQDDGRGLVLPGYKYLPGFNGLDKGE 63
Db	3 DGYLPDWLEDNLSGIREWDLKPAPKPKANQQKQDDGRGLVLPGYKYLPGFNGLDKGE 62
QY	64 PVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLQEDTSFGNLRGAVFQAKK 123
Db	63 PVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLQEDTSFGNLRGAVFQAKK 122
QY	124 RVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNF-QQTGDSSEV 182
Db	123 RVLEPLGLVEEGAKTAPGKKRPLIESPQPDSSSTGIGKKGQPAKKLVFEDETGAGDGP 182
QY	183 PDQPLGEPATPAAGVPTTMSAGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVIT 242
Db	183 PEGSTSG-----AMSDDSEMAAGAAVEGGGADGVGNASGDWHCDSTWSEGHVIT 236
QY	243 STRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYFDENRFHCHFSRPDWRQLIN 302
Db	237 STRTWALPTYNNHLYKRLGE---SLQSNVTYNGFSTPWGYFDENRFHCHFSRPDWRQLIN 292
QY	303 NNWGFRRPKRLNFKLFIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGC 362
Db	293 NNWGMRRPKAMRVKIFNIQVKEVTTNGETTVANNLTSTVQIFADSSYELPYVMDAQEGS 352
QY	363 LPPFPADVFMIPQYGY--LTLNNGSQAVGRSSPYCLEYFPFSQMLRTGNFTFSYTPEEV 419
Db	353 LPPFPNDVFMVPQYGYCGLVTGNTSQQQTDNRNAYCYCLEYFPFSQMLRTGNFTFSYFEKV 412
QY	420 PFHSSYAHSSQSLDRLMNPLIDQYLYLNRNTQ---QSGSAQNKOLLFSRGPAGMSVQP 475
Db	413 PFHSMYAHSSQSLDRLMNPLIDQYLGWLGQSTTTGTTLNAGTATTN---FTKLRPTNFSNPK 469
QY	476 KNWLPGPCYRQQRVSKTKTDNNNSNFTWTGAS--KY---NLNGRESIINPGTAMASHK 528
Db	470 KNWLPGPSIKQQGFSKTA--NQNYKIPATGSDSLIKYETHSTLDGRWSALTGPPPMATAG 527
QY	529 DDEDKFFPMGSGVMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFSQS 588
Db	528 PADSK-FSNSQLIFAGPKQNGTATVPGLTIFSEEELAAATNATDMDWGNLPGGDQSNS 586
QY	589 TDPATGDVHAMGALPGMWQDRDVYLGQPIWAKIPHTDGHFHPSPMLGGFGLKNPPQIL 648
Db	587 NLPTVDRLTALGAVPGMWQNRDIYYQGPWAKIPHTDGHFHPSPGLGGFGLKHPPQIF 646
QY	649 IKNTPVPANPPAEFSATKFASFITQYSTGQVSVEIEWELQENSKRWNPVQYTSNYAKS 708
Db	647 IKNTPVPANPATTFSSTPVNSFITQYSTGQVSQVQIDWEIQERSKRWNPVQYTSNYGQQ 706
QY	709 ANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
Db	707 NSLLWAPDAAGKYTEPRAIGTRYLTHHL 734
RESULT 10	
O92917	PRELIMINARY; PRT; 533 AA.
ID	O92917
AC	O92917;
DT	01-NOV-1998 (TrEMBLrel. 08, Created)
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	Major coat protein VP3.
OS	Adeno-associated virus 2 (AAV2).
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX	NCBI_TaxID=10804;
RN	[1]
RP	SEQUENCE FROM N.A.


```
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
RT proteins affect viral infectivity: lack of an RGD integrin-binding
RT motif.";
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03779.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
FT VARIANT 351 351 D -> N.
FT VARIANT 365 365 T -> N.
FT VARIANT 475 476 QV -> HV.
FT VARIANT 508 508 V -> R.
SQ SEQUENCE 533 AA; 60063 MW; 9E4D8BC25810D4F0 CRC64;

Query Match 62.2%; Score 2481.5; DB 12; Length 533;
Best Local Similarity 83.3%; Pred. No. 1e-162;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 203 MASGGAPMADNNEGADGVGNAGNWHCDSTWLGLDRVITSTRTWALPTYNNHLYKQISS 262
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MATSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITSTRTWALPTYNNHLYKQISS 60

QY 263 ASTGASNDNHFGYSTPWGYFDNRFHCFSPRDWQRLNNNWGFRPKRLNFKLFNIQVK 322
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 QS-GASNDNHFGYSTPWGYFDNRFHCFSPRDWQRLNNNWGFRPKRLNFKLFNIQVK 119

QY 323 EVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPPPADVFMIPOYGYLTIN 382
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
120 EVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGLPPPPADVFVMPQYGYLTIN 179

QY 383 NGSQAVGRSSFYCLEYFPQSMLRTGNNTFSYTFEEVPEFSSYAHQSQSLDRLMNLIDQY 442
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
180 NGSQAVGRSSFYCLEYFPQSMLRTGNNTFSYTFEDVPEFSSYAHQSQSLDRLMNLIDQY 239

QY 443 LYILNRTONQSGSAQNKKLLFRSGSPAGMSVQPKNLPGPCYRQQRVSKTKTDNNNSFT 502
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
240 LYILSRNTNTPSGTTTQSRQLQFSQAGASDIRDQSRNLWLPGPCYRQQRVSKTSADNNNSEYS 299

QY 503 WTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMMSGVMIFGKESAGASNTALDNVMTD 562
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
300 WTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVLIFPGQSEKTNVDIEKVMITD 359

QY 563 EBEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQDRDVYLGQPIWAKI 622
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
360 EBEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLGQPIWAKI 419

QY 623 PHTDGHFHPSPLMGGFGLKNPPQIILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 682
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
420 PHTDGHFHPSPLMGGFGLKHPPQIILIKNTVPANPSTTSAAKFAFITQYSTGQVSVE 479

QY 683 IEWELQENSKRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
480 IEWELQENSKRWNPYQYTSNYKSNVNVDFTVDTNGVYSEPRPIGTRYLTRNL 533

RESULT 11
Q9YIJ1 PRELIMINARY; PRT; 724 AA.
AC Q9YIJ1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Capsid protein.
GN CAP OR VP1.
OS Adeno-associated virus 5.
```

```
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=82300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098980; PubMed=9882294;
RA Bantel-Schaal U., Delius H., Schmidt R., Zur Hausen H.;
RT "Human adeno-associated virus type 5 is only distantly related to
RT other known primate helper-dependent parvoviruses.";
RL J. Virol. 73:939-947(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99099022; PubMed=9882336;
RA Chiorini J.A., Kim F., Yang L., Kotin R.M.;
RT "Cloning and characterization of adeno-associated virus type 5.";
RL J. Virol. 73:1309-1319(1999).
DR EMBL; Y18065; CAA77024.1; -.
DR EMBL; AF085716; AAD13756.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 724 AA; 80424 MW; BBC332B88258D34C CRC64;

Query Match 55.7%; Score 2220; DB 12; Length 724;
Best Local Similarity 58.6%; Pred. No. 1.7e-144;
Matches 431; Conservative 81; Mismatches 198; Indels 26; Gaps 9;

QY 8 PDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPYKYKLPFNGLDKGEPVNA 67
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
8 PDWLEE-VGEGLEFLGLEAGPPKPKPNQQHQDQARGLVLPYNYLPGNGLDRGEPVNR 66

QY 68 AAAAALEHDKAYDQQLKAGDNPYLRYNHADAERFQERLQEDTSFGNGLGRAVFAQKRVLE 127
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 ADEVAREHDISYNEQLEAGDNPYLKYNHADAERFQEKLADDTSFNGNLGKAVFAQKRVLE 126

QY 128 PLGLVEEGAKTAPGKKRPVEQSPQEPDSSGIGTKGQPAKRLNFGQTGDSESVDPDQP 187
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
127 PFGLVEEGAKTAPGKRIDDHFPKRKKART-----EEDSKP-----STSSDABAGPSGSQ 176

QY 188 LGEPPATPA-AVGPTTMASGGGAPMADNNEGADGVGNAGNWHCDSTWLGLDRVITSTRT 246
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
177 QLQIPAQPASSLGADTMSAGGGPLGDNNOQADGVGNASGDWHCDSTWMGDRVVTKSTRT 236

QY 247 WALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFHCFSPRDWQRLNNNWG 306
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 WVLPSYNNHQYREIKSGSVDSGNANAYFGYSTPWGYFDNRFHSHWSPRDWQRLNNYWG 296

QY 307 FRPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPF 366
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
297 FRPRSLRVKIFNIQVKEVTVQDSTTTIANNLSTVQVFTDDDYQLPYVVGNGTEGCLPAF 356

QY 367 PADVFMIPQYGYLTIN--NGSQAVGRSSFYCLEYFPQSMLRTGNNTFSPYTFEEVPPHSS 424
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
357 PPQVFTLPQYGYATLNRDNTENPTERSFFCLEYFPKMLRTGNNTFNTFEEVPPHSS 416

QY 425 YAHSQSLDRLMNLIDQYLYLNRTONQSGSAQNKKLLFRSGSPAGMSVQP-KNWLPGPC 483
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
417 FAPSQNLFLKANPLVDQYLYRFVSTNNTGGVQFNKNL-----AGRYANTYKNWFPGPM 469

QY 484 YRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMMSGVMIF 543
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
470 GRTQGNLGSVNRASVSFAFATNRMELEGASYQVPPQPNGMTNNLQGSNTVALENTMIF 529

QY 544 GKESAGASNTAL---DNVMITDEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG 600
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
530 NSQAPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSSTTAPATGTYNLQE 589

QY 601 ALPGMVWQDRDVYLGQPIWAKIPHTDGHFHPSPLMGGFGLKNPPQIILIKNTVPANPPA 660
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
590 IVPGSVMWMERDVYLGQPIWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVPGN-IT 648

QY 661 EFSATKFAFITQYSTGQVSVEIEWELQENSKRWNPVEQYTSNYAKSANVDFTVDNNGL 720
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
649 SFSDVPVSSFITQYSTGQVTMEMEWELKSKRWNPYQYTSNYNDPQVDFADPDSTGE 708
```



```
Qy 721 YTEPRPIGTRYLTRPL 736
Db 709 YRTTRPIGTRYLTRPL 724

RESULT 12
Q83290
ID Q83290 PRELIMINARY; PRT; 732 AA.
AC Q83290;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Muscovy duck parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=37325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=89384/France;
RX MEDLINE=96406928; PubMed=8811015;
RA Le Gall-Recule G., Jestin V., Chagnaud P., Blanchard P., Jestin A.;
RT "Expression of muscovy duck parvovirus capsid proteins (VP2 and VP3)
in a baculovirus expression system and demonstration of immunity
induced by the recombinant protein.";
RL J. Gen. Virol. 77:2159-2163(1996).
DR EMBL; Z68272; CAA92575.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat;_1.
FT CHAIN 146 732 VP2 CAPSID PROTEIN.
FT CHAIN 199 732 VP3 CAPSID PROTEIN.
SQ SEQUENCE 732 AA; 81364 MW; DE70DCCAB215F4E2 CRC64;

Query Match 54.9%; Score 2190.5; DB 12; Length 732;
Best Local Similarity 55.7%; Pred. No. 1.9e-142;
Matches 422; Conservative 88; Mismatches 184; Indels 63; Gaps 13;

Qy 9 DWLEDNLSEGIREWDLKPGAPKPKANQQKQ-----DDGRGLVLPGYKYLGPFN 57
Db 10 DWYE-----TAAASWRHLKAGAPKPKSNQQSQSVSTRDKPQKDNRRGFVLPGYKYLPGN 65

Qy 58 GLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERLQEDTSFGGNLGRA 117
Db 66 GLDKGPPVNKADSVALEHDKAYDQQLKAGDNPYIKFNHADQEFIDNLQGDTSFGGNLGKA 125

Qy 118 VFQAKKRVLPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQQQPAKKRLNFGQTG 177
Db 126 VFQAKKRILEPLGLVEEPVNTAPAKK-----SSG-KLTDHYPIVKKPKLSE-- 170

Qy 178 DSESVDPPQPLG-----EPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASG 226
Db 171 --ENSPPSNSGGEASAAATEGSEPVAA-----NMAEGSGGAMGDSAGGADGVGNASG 222

Qy 227 NMHCSTWLGDVRVITSTRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWGYPDFN 286
Db 223 NMHCDSQWLGDVTIKTTRTWLPSYNNHIYKAITSGTNPDSN-TQYAGYSTPWGYPDFN 281

Qy 287 RFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFS 346
Db 282 RFHCHFSPRDWQRLINNNHWGIRPKALKFKIFNVQVKEVTTQDQTKTIANNLTSTIQIF 341

Qy 347 SEYQLPYVLGSAHQGCLPPFPADVFMIPQGYLTNLN---NGSQAVGRSSFYCLEYFPSQM 403
Db 342 NEHQLPYVLGSATEGTMPPFPSPDVVALPQYGYCTMHTNQSGARFNDRSFYCLEYFPSQM 401

Qy 404 LRTGNNTFSYTFEEVPFHSSYAHQSQSLDRMLNPLIDQYLYLNRNTQNQSGSAQNKDLF 463
Db 402 LRTGNNPFSEFEFEEVPFHSMFAHSQDLDRMLNPLLDQYLNWFSEV-NGGRNAQ-----F 455

Qy 464 SRGSPAGMSVQPKNWLPGPCYRQQRVSKTK--TDNNNSNFTWTGASKYNLNGRESINPG 521
Db 456 KKA VKGAFGAMGRNWLPGPKLLDQVRAYSGGTDNYANWSIWSKGNKVLKOREYLLQPG 515

Qy 522 TAMASHKODEDKFFPMSGVMIFGKE--SAGASNTALDNVMTTDEEIKATNPVATERFCT 579
```

```
Db 516 PVATTHTEQASSVPAQNIIGIAKDPYRSGSTLAGISDIMVTDEQEIAPTNGVGWRPYGL 575
Qy 580 VAVNFQSSSTDPATGDVHAMGALPGVMWQDRDVYLOGPIWAKIPHTDGHFHPSPLMGGFG 639
Db 576 TVTNEQNTTAPTNAEVLGALPGVMWQNRDIYLOGPIWAKIPKTDGKFPSPNLGGFG 635

Qy 640 LKNPPPOILIKNTVPANPPAEFSATKPFASFIQYSTGQVSVEIEWELQKNSKRWNPV 699
Db 636 LHNPPPOVFIKNTVPADPPLEYVQKWNYSYITQYSTGQCTVEMVWELKNSKRWNP 695

Qy 700 QYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
Db 696 QFTSNFGNRTSTMFAPNETGGYVEDRLIGTRYLTQNL 732

RESULT 13
Q67666
ID Q67666 PRELIMINARY; PRT; 732 AA.
AC Q67666;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VP1.
GN VP1.
OS Goose parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=38251;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Virulent B;
RX MEDLINE=96010229; PubMed=7571426;
RA Zadori Z., Stefancsik R., Rauch T., Kisaary J.;
RT "Analysis of the complete nucleotide sequences of goose and muscovy
duck parvoviruses indicates common ancestral origin with adeno-
virus associated virus 2.";
RL Virology 212:562-573(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Virulent B;
RA Zadori Z.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U25749; AAA83230.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat;_1.
SQ SEQUENCE 732 AA; 81340 MW; 339507C61D47B52C CRC64;

Query Match 54.5%; Score 2173.5; DB 12; Length 732;
Best Local Similarity 54.7%; Pred. No. 2.8e-141;
Matches 414; Conservative 108; Mismatches 176; Indels 59; Gaps 14;

Qy 9 DWLEDNLSEGIREWDLKPGAPKPKANQQKQ-----DDGRGLVLPGYKYLGPFN 57
Db 6 DSFEWYETAAASWRNLKAGAPQPKPNQSQSVSPDREPERKDNRRGFVLPGYKYLPGN 65

Qy 58 GLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERLQEDTSFGGNLGRA 117
Db 66 GLDKGPPVNKADSVALEHDKAYDQQLKAGDNPYIKFNHADQEFIDSLQDDQSFSGNLGKA 125

Qy 118 VFQAKKRVLPLGLVEEGAKTAPGKKR-----PVEQSPQ-EPDSSSGIGKTQQQPAK 168
Db 126 VFQAKKRILEPFGVLVEDPVNTAPAKKNTKLTLDHYPVVKKPLTEEVSAAGSSAVQDG- 184

Qy 169 KRLNFGQTGDSESVDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNW 228
Db 185 -----GATAE-----GTEP-----VAASEMAEGGGAMGDSGGGADGVGNASGNW 224

Qy 229 HCDSTWLGDVRVITSTRTWALPTYNHLYKQISSASTGASNDNH--YFGYSTPWGYPDFN 286
Db 225 HCDSQWMGNTVITKTRTWLPSYNNHIYKAITS---GTSQDANVQYAGYSTPWGYPDFN 281

Qy 287 RFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFS 346
```

Db 282 RFCHFSPRDWORLINNHWGIRPKSLKFIENVQKEVTTQDQTKTIANNLTSTIQVFTD 341
QY 347 SEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTLN---NGSQAVGRSSFYCYLEYFPSQM 403
Db 342 DEHQLPYVLGSATEGTMPPFPSPDVYALPOYGYCTMHTNQNGARENDRSAFYCYLEYFPSQM 401
QY 404 LRTGNFTFSYTFEEVPHSSYAHSQSLDRLMNPPLIDQYLYLNRNTQNSGSAQNKDLLF 463
Db 402 LRTGNNFETFDVEVPFHSMFAHSQDLRLMNPPLVDQYLNWFNEV-DSSRNAQ-----F 455
QY 464 SRGSPAGMSVQPKNWLPGPCYRQRV-SKTKTNNNSNFT-WTGASKYNLNGRESIINPG 521
Db 456 KKAVKGAYGTMRNWLPGPKFLDQVRAYTGGTDNYANWNIWSNGKNVNLKDRQYLLQPG 515
QY 522 TAMASHKDDDEKFFPMSGVMIFGKE--SAGASNTALDNVMITDEEEIKATNPVATERFGT 579
Db 516 PVSATYTEGEASSLPAQNILGIAKDPYRSGSTTAGISDIMVTEQEVAPTNGVGWKKPYGR 575
QY 580 VAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLGQPIWAKIPHTDGHFHPSPLMGGFG 639
Db 576 TVTNEQNTTAPTSSDLVDLALPGMVWQNRDIYLGQPIGAKIPKTDGKFHPSPLGGFG 635
QY 640 LKNPPPPQILIKNTVPANPPAEPFSAFKFASFITQYSTGQSVSEIEWELQKENSKRWNPEV 699
Db 636 LHNPPPPQVFIKNTVPADPPVEYVHOKWNSYITQYSTGQCTVEMVWELRKENSKRWNPEI 695
QY 700 QYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 736
Db 696 QFTSNFSNRTSIMFAPNETGGYVEDRLIGTRYLTQNL 732

RESULT 14
Q83289 PRELIMINARY; PRT; 732 AA.
AC Q83289;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VP1 protein.
GN VP1.
OS Muscovy duck parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=37325;
RN [1]
RP SEQUENCE FROM N.A.
RA Zadori Z., Erdei J., Nagy J., Kisari J.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X75093; CAA52984.1; --
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 732 AA; 81284 MW; 9FF6955BC15DF3AA CRC64;

Query Match 54.4%; Score 2168.5; DB 12; Length 732;
Best Local Similarity 55.2%; Pred. No. 6.1e-141;
Matches 418; Conservative 90; Mismatches 186; Indels 63; Gaps 13;
QY 9 DWLEDNLSEGIREWMDLKPAPKPKANQQKQ-----DDGRGLVLPGYKYLGPFN 57
Db 10 DWYE----TAAASWRHLKAGAPKPKSNQQSVSTDRKPKQRKNNRGFVLPGYKYVGPGN 65
QY 58 GLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFFQERLQEDTSFGGNLGRA 117
Db 66 GLDKGPPVNAKADSVALEHDKAYDQQLKAGDNPYIKFKHADQBFIDNLQGDTSFGGNLGA 125
QY 118 VFQAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGICKTGQQPAKRLNFGGTG 177
Db 126 VFQAKKRILEPLGLVEEPVNTAPAK-----SSG-KLTDHDPVKKPKLSE-- 170
QY 178 DSEVPDPQPLG-----EPPATPAAVGPTTMASGGGAPMADNNEGADVGNASG 226
Db 171 --ENSPPSPNSNGGEASAAATEGSEPVAA-----NMAEGSGAMGDSAGGADVGNASG 222
QY 227 NWHCDSTWLGDVRVITSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDEN 286

Db 223 NWHCDSQWLGDTVITKTTRTWLPSYNNHMYQAITSCTNPDSN-TQYAGYSTPWGYFDEN 281
QY 287 RFCHFSPRDWORLINNHWGFRPKRLNFKLFENIQVKEVTTNDGVTTIANNLSTVQVPSD 346
Db 282 RFCHFSPRDWORLINNHWGIRPKALKFKFIENVQKEVTTQDQTKTIANNLTSTIQIFTD 341
QY 347 SEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTLN---NGSQAVGRSSFYCYLEYFPSQM 403
Db 342 NEHQLPYVLGSATEGTMPPFPSPDVYALPOYGYCTMHTNQSGARFNDRSAFYCYLEYFPSQM 401
QY 404 LRTGNFTFSYTFEEVPHSSYAHSQSLDRLMNPPLIDQYLYLNRNTQNSGSAQNKDLLF 463
Db 402 LRTGNNFESFEFEEVPFHSMFAHSQDLRLMNPPLDQYLNWFNEV-NGGRNAQ-----F 455
QY 464 SRGSPAGMSVQPKNWLPGPCYRQRVSKTK--TDNNNSNFTWTGASKYNLNGRESIINPG 521
Db 456 KKAVKGAFGAMGRNWLPGPKLLDQVRAYSGGTDNYANWSIWSKGNKVFLKDRYLLQPG 515
QY 522 TAMASHKDDDEKFFPMSGVMIFGKE--SAGASNTALDNVMITDEEEIKATNPVATERFGT 579
Db 516 PVATHTTEDQASSVPAQNIIGIAKDPYRSGSTLAGISDIMVTEQEIAPTNGVGRPYGL 575
QY 580 VAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLGQPIWAKIPHTDGHFHPSPLMGGFG 639
Db 576 TVTNEQNTTAPTNAELEVLGALPGMVWQNRDIYLGQPIWAKIPKTDGKPHPSPLGGFG 635
QY 640 LKNPPPPQILIKNTVPANPPAEPFSAFKFASFITQYSTGQSVSEIEWELQKENSKRWNPEV 699
Db 636 LHNPPPPQVFIKNTVPADPPLEYVHOKWNSYITQYSTGQCTVEMVWELRKENSKRWNPEI 695
QY 700 QYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 736
Db 696 QFTSNFGNRTSTMFAPNETGGYVEDRLIGTRYLTQNL 732

RESULT 15
Q8V395 PRELIMINARY; PRT; 732 AA.
AC Q8V395;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Capsid protein VP.
GN VP.
OS Goose parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=38251;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPV-YG;
RA Ge Y., You Y., Xu Q.;
RT "Analysis of the major open reading frames' nucleotide sequences in
RT Goose parvovirus GPV-YG strain isolated in China."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416726; AAL37722.1; --
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 732 AA; 81456 MW; 73F2E4BC769744B6 CRC64;

Query Match 54.3%; Score 2165.5; DB 12; Length 732;
Best Local Similarity 54.8%; Pred. No. 9.9e-141;
Matches 415; Conservative 101; Mismatches 182; Indels 59; Gaps 14;
QY 9 DWLEDNLSEGIREWMDLKPAPKPKANQQKQ-----DDGRGLVLPGYKYLGPFN 57
Db 6 DSFEWYETAASWRNLKAGAPHPKPNQQTQSVSPAREPERRDSNRGFVLPGYKYLPGN 65
QY 58 GLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFFQERLQEDTSFGGNLGRA 117
Db 66 GLDKGPPVNAKADSVALEHDKAYDQQLKAGDNPYIKFNHADQDFIDSLQDDRSFGGNLGA 125
QY 118 VFQAKKRVLEPLGLVEEGAKTAPGKKR-----PVEQSPQ-EPDSSSGICKTGQQPAK 168

Db 126 VFQAKRILEPFLVEDSINTAPVKKNTGKLTDPVVKPKLTFEEVSAGGTSVQDG- 184
Qy 169 KRLNFGQTGDSESDPQPLGEPPATPAAVGPTTMASGGAPMADNNEGADGVGNASGNW 228
Db 185 -----GATAE-----GTEP-----VAASEMAEGGGALGDASGGADGVGNASGNW 224
Qy 229 HCDSTWLGDRVITSTRTWTALPTYNNHLYKQISSASTGASNDN--HYFGYSTPWGYFDFN 286
Db 225 HCDSQWMGNTVITKTRTWLWLPSSNNHIYKAITS---GTSQDATVQYAGYSTPWGYFDFN 281
Qy 287 RFHCHFSRDPWQRLINNNWGRFPKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVFS 346
Db 282 RFHCHFSRDPWQRLINNNHWGIRPKSLKFKIFNVQVKEVTTQDTKTIANNLTSTIQVFTD 341
Qy 347 SEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTLN---NGSOAVGRSSFYCYLFPPSQM 403
Db 342 DEHQLPYVLGSATEGTMPPFPSPDVYALPQYGYCTMHTNQNGARENDRSAFYCYLFPPSQM 401
Qy 404 LRTGNNFTFSYTFEEVPHSSYAHQSLSLDRMLNPLIDQYLYLNRNTQNSGSAQNKDLLF 463
Db 402 LRTGNNFEFTDFEEVPHSMFAHSQDLDRMLNPLVDQYLNWFNEV-DSSRKAQ-----F 455
Qy 464 SRGSPAGMSVQPKNWLPGPCYRQORV-SKTKTDNNNSNFT-WTGASKYNLNGRESIINPG 521
Db 456 KKA VKGAYGTMGRNWLPGPKLLDQVRAYTGGTDNYANWNINWSNGKNVNLKDRQYLLQPG 515
Qy 522 TAMASHKODEDKFFPMSGVNIFGKE--SAGASNTALDNVMITDEEEIKATNPVATERFGT 579
Db 516 PVSATHTKVEASSIPAQNILGLAKOPYRSGSTTAGISDIMVTDEQEVAPTNGVGWKPYGK 575
Qy 580 VAVNFOSSSTDPA TGDVHAMGALPGMVWQDRD VYLGQPIWAKIPHTDGHFHPSPLMGGFG 639
Db 576 TVTNEQNTTAPTSSDLDVLGALPGMVWQNRDIYLGQPIWAKIPKTDGKFHPSPNLGGFG 635
Qy 640 LKNPPPPQILIKNTVPANPPAFSATAKFASFITQYSTGVSVVEIEWELQENSKRWNP 699
Db 636 LKNPPPPQVFIKNTVPADPPVVEYVHQWNSYITQYSSGQCTVEMVWELRKENSKRWNP 695
Qy 700 QYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
Db 696 QYTSNPSDRTSIMFAPNETGGYIEDRLIGTRYLTQNL 732

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:04:27 ; Search time 33.0787 Seconds
(without alignments)
4549.706 Million cell updates/sec

Title: US-09-807-802A-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3989	100.0	736	12 US-10-291-583-64	Sequence 64, Appl
2	3989	100.0	736	12 US-10-423-704A-5	Sequence 5, Appli
3	3963	99.3	736	12 US-10-291-583-65	Sequence 65, Appl
4	3494	87.6	736	12 US-10-291-583-71	Sequence 71, Appl
5	3494	87.6	736	12 US-10-423-704A-6	Sequence 6, Appli
6	3473	87.1	736	12 US-10-291-583-97	Sequence 97, Appl
7	3467	86.9	736	12 US-10-291-583-100	Sequence 100, App
8	3467	86.9	736	12 US-10-423-704A-7	Sequence 7, Appli
9	3464	86.8	736	12 US-10-291-583-96	Sequence 96, Appl
10	3460	86.7	736	12 US-10-291-583-98	Sequence 98, Appl
11	3457	86.7	736	12 US-10-291-583-99	Sequence 99, Appl
12	3453	86.6	738	12 US-10-291-583-85	Sequence 85, Appl
13	3448	86.4	738	12 US-10-291-583-79	Sequence 79, Appl
14	3448	86.4	738	12 US-10-291-583-81	Sequence 81, Appl
15	3445	86.4	738	12 US-10-291-583-93	Sequence 93, Appl

16	3445	86.4	738	12	US-10-291-583-94	Sequence 94, Appl
17	3442.5	86.3	737	12	US-10-291-583-2	Sequence 2, Appli
18	3442.5	86.3	737	12	US-10-423-704A-8	Sequence 8, Appli
19	3441	86.3	738	12	US-10-291-583-80	Sequence 80, Appl
20	3440	86.2	738	12	US-10-291-583-84	Sequence 84, Appl
21	3438	86.2	738	12	US-10-291-583-92	Sequence 92, Appl
22	3434	86.1	728	12	US-10-291-583-108	Sequence 108, App
23	3432	86.0	738	12	US-10-291-583-91	Sequence 91, Appl
24	3429	86.0	738	12	US-10-291-583-82	Sequence 82, Appl
25	3424.5	85.8	737	12	US-10-291-583-72	Sequence 72, Appl
26	3422.5	85.8	729	12	US-10-291-583-110	Sequence 110, App
27	3422	85.8	738	12	US-10-291-583-83	Sequence 83, Appl
28	3420.5	85.7	733	12	US-10-291-583-86	Sequence 86, Appl
29	3420	85.7	728	12	US-10-291-583-102	Sequence 102, App
30	3417.5	85.7	729	12	US-10-291-583-111	Sequence 111, App
31	3416.5	85.6	733	12	US-10-291-583-87	Sequence 87, Appl
32	3415.5	85.6	735	12	US-10-291-583-67	Sequence 67, Appl
33	3415	85.6	728	12	US-10-291-583-106	Sequence 106, App
34	3414.5	85.6	729	12	US-10-291-583-109	Sequence 109, App
35	3414.5	85.6	731	12	US-10-291-583-89	Sequence 89, Appl
36	3413.5	85.6	735	12	US-10-291-583-68	Sequence 68, Appl
37	3413.5	85.6	735	12	US-10-291-583-112	Sequence 112, App
38	3413	85.6	728	12	US-10-291-583-107	Sequence 107, App
39	3411.5	85.5	735	12	US-10-291-583-69	Sequence 69, Appl
40	3408.5	85.4	731	12	US-10-291-583-88	Sequence 88, Appl
41	3407.5	85.4	735	12	US-10-291-583-66	Sequence 66, Appl
42	3406	85.4	728	12	US-10-291-583-105	Sequence 105, App
43	3405	85.4	738	12	US-10-291-583-95	Sequence 95, Appl
44	3405	85.4	738	12	US-10-423-704A-2	Sequence 2, Appli
45	3402.5	85.3	735	12	US-10-291-583-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1

US-10-291-583-64
; Sequence 64, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV1
US-10-291-583-64

Query Match 100.0%; Score 3989; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDNLSEGREWDLKPGAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60
|||||

Db 1 MAADGYLPDWLEDNLSEGREWDLKPGAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60
|||||

QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEEFQERLQEDTSFGNGLGRAVFQ 120

Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
QY 181 SVPDPQPLGEPPTAAAVGPTTASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAAVGPTTASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRL 300
QY 301 INNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
Db 301 INNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
Db 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
QY 421 FHSSYAHQSLSLDRLMNPLIDQLYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLP 480
Db 421 FHSSYAHQSLSLDRLMNPLIDQLYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEDKFFPMSGV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEDKFFPMSGV 540
QY 541 MIFGESAGASNTALDNNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
Db 541 MIFGESAGASNTALDNNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
QY 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPV PANPPA 660
Db 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPV PANPPA 660
QY 661 EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 720
Db 661 EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736
RESULT 2
US-10-423-704A-5
; Sequence 5, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 736

; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-423-704A-5
Query Match 100.0%; Score 3989; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLPDWLEDNLSEGIREWWDLPGAPKPKANQQKQDDGRGLVLPGYKYLGPFPNGLD 60
Db 1 MAADGYLPDWLEDNLSEGIREWWDLPGAPKPKANQQKQDDGRGLVLPGYKYLGPFPNGLD 60
QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQ 120
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
QY 181 SVPDPQPLGEPPTAAAVGPTTASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAAVGPTTASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRL 300
QY 301 INNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
Db 301 INNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
Db 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
QY 421 FHSSYAHQSLSLDRLMNPLIDQLYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLP 480
Db 421 FHSSYAHQSLSLDRLMNPLIDQLYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEDKFFPMSGV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEDKFFPMSGV 540
QY 541 MIFGESAGASNTALDNNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
Db 541 MIFGESAGASNTALDNNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
QY 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPV PANPPA 660
Db 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPV PANPPA 660
QY 661 EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 720
Db 661 EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736
RESULT 3
US-10-291-583-65
; Sequence 65, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (A
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583

```

; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV6VP1
US-10-291-583-65
```

```

Query Match          99.3%; Score 3963; DB 12; Length 736;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```

Qy 1 MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPENGLD 60
Db 1 MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPENGLD 60

Qy 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERFQRLQEDTSFGGNLGRAVFQ 120
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERFQRLQEDTSFGGNLGRAVFQ 120

Qy 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180

Qy 181 SVPDPQPLGEPPTAAVGGTPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAVGGTPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240

Qy 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRL 300

Qy 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
Db 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360

Qy 361 GCLPPFPADVPMIPOYGylTLNNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGylTLNNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEDVP 420

Qy 421 FHSSYAHQSLSLDRLMNLPLIDQLYLYLNRNQSGSAQNKDLSFSGSPAGMSVQPKNWL 480
Db 421 FHSSYAHQSLSLDRLMNLPLIDQLYLYLNRNQSGSAQNKDLSFSGSPAGMSVQPKNWL 480

Qy 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGV 540
Db 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGV 540

Qy 541 MIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPRATGDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPRATGDVHAMG 600

Qy 601 ALPGMWQDRDQVYLOGPIWAKIPHDTGHFHPSPLMGGFGLKNPPPPQILIKNTVPVANPPA 660
Db 601 ALPGMWQDRDQVYLOGPIWAKIPHDTGHFHPSPLMGGFGLKNPPPPQILIKNTVPVANPPA 660

Qy 661 EFSATKFAFITQYSTGQVSVEIEWELQKENSQRWNPEVQTSNYAKSANVDFTVDNNGL 720
Db 661 EFSATKFAFITQYSTGQVSVEIEWELQKENSQRWNPEVQTSNYAKSANVDFTVDNNGL 720

Qy 721 YTEPRFIGTRYLTRPL 736
Db 721 YTEPRFIGTRYLTRPL 736
```

```

RESULT 4
US-10-291-583-71
; Sequence 71, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV3
US-10-291-583-71
```

```

Query Match          87.6%; Score 3494; DB 12; Length 736;
Best Local Similarity 86.4%; Pred. No. 2.e-295;
Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;
```

```

Qy 1 MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPENGLD 60
Db 1 MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPENGLD 60

Qy 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERFQRLQEDTSFGGNLGRAVFQ 120
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERFQRLQEDTSFGGNLGRAVFQ 120

Qy 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180

Qy 181 SVPDPQPLGEPPTAAVGGTPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAVGGTPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240

Qy 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRL 300

Qy 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
Db 301 INNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360

Qy 361 GCLPPFPADVPMIPOYGylTLNNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGylTLNNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEDVP 419

Qy 421 FHSSYAHQSLSLDRLMNLPLIDQLYLYLNRNQSGSAQNKDLSFSGSPAGMSVQPKNWL 479
Db 421 FHSSYAHQSLSLDRLMNLPLIDQLYLYLNRNQSGSAQNKDLSFSGSPAGMSVQPKNWL 479

Qy 480 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db 480 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539

Qy 540 VMI FGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPRATGDVHAM 599
Db 540 NLI FGKEGTTASNAELDNVMITDEEEIRTNVPATEQYGTVANNLQSSNTAPTGTGVNHQ 599
```


Db 181 SVPDPQPLGEPAPASGLGPNITWASGGAPMADNNEGADGVGNSSGNWHCDSTWLGDRVI 240
Qy 241 TTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDNRFHCHFSPRDWQR 299
Db 241 TTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDNRFHCHFSPRDWQR 300
Qy 300 LINNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDESEYQLPYVLGSAH 359
Db 301 LINNWGFRPKRLNFKLFNIQVKEVTTNEGTKTIANNLTSTVQVFTDSEYQLPYVLGSAH 360
Qy 360 QGCLPPFPADVFMIPQYGYLTLLNNGSQAVGRSSFYCYLEYFPPSQMLRTGNNFTFSYTFEEV 419
Db 361 QGCLPPFPADVFMVPQYGYLTLLNNGSQALGRSSFYCYLEYFPPSQMLRTGNNFQFSYTFEDV 420
Qy 420 PFHSSYAHQSLSLDRMLNPLIDQYLYLNRNTQNSGAQNKDLLFSRGSPAGMSVQPKNWL 479
Db 421 PFHSSYAHQSLSLDRMLNPLIDQYLYLVRTQT-TGTGGTQTALFSAQAGPSSMANQARNWV 479
Qy 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db 480 PGPCYRQORVSTTTNQNNNSNFAWTGAAKFKLNGRDSLMPNPGVAMASHKDDDRFFPSSG 539
Qy 540 VMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAM 599
Db 540 VLIFGKQAGNDGVDYSQVLITDEEEIKATNPVATEEYGAVAINNQAANTQAOTGLVHNQ 599
Qy 600 GALPGMWQDRDQVYLGQPIWAKIPHTDGHFHPSPMLMGFGFLGNPPQILIKNTVPANPP 659
Db 600 GVIPGMVWQNRDQVYLGQPIWAKIPHTDGNFHPSPMLMGFGFLGNPPQILIKNTVPADPP 659
Qy 660 AEFSATKPFASITQYSTQGVSVSVEIEWELQENSKRWNPVQYTSNYAKSANVDFTVDNNG 719
Db 660 LTFNQAKLNSFITQYSTQGVSVSVEIEWELQENSKRWNPVQYTSNYAKSTNVDFAVNTEG 719
Qy 720 LYTEPRPIGTRYLTRPL 736
Db 720 VYSEPRPIGTRYLTRNL 736

RESULT 7
US-10-291-583-100
; Sequence 100, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV9
US-10-291-583-100

Query Match 86.9%; Score 3467; DB 12; Length 736;
Best Local Similarity 85.5%; Pred. No. 4.9e-293;
Matches 630; Conservative 44; Mismatches 61; Indels 2; Gaps 2;

Qy 1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGEFVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFO 120
Db 61 KGEFVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFO 120
Qy 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGTKTQQPAKQRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGTKSQQPAKQRLNFGQTGDSE 180
Qy 181 SVPDPQPLGEPATPAAVGTPTMASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPEAPSGLGPNITWASGGAPMADNNEGADGVGNSSGNWHCDSTWLGDRVI 240
Qy 241 TTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDNRFHCHFSPRDWQR 299
Db 241 TTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDNRFHCHFSPRDWQR 300
Qy 300 LINNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDESEYQLPYVLGSAH 359
Db 301 LINNWGFRPKRLNFKLFNIQVKEVTTNEGTKTIANNLTSTVQVFTDSEYQLPYVLGSAH 360
Qy 360 QGCLPPFPADVFMIPQYGYLTLLNNGSQAVGRSSFYCYLEYFPPSQMLRTGNNFTFSYTFEEV 419
Db 361 QGCLPPFPADVFMVPQYGYLTLLNNGSQALGRSSFYCYLEYFPPSQMLRTGNNFQFSYTFEDV 420
Qy 420 PFHSSYAHQSLSLDRMLNPLIDQYLYLNRNTQNSGAQNKDLLFSRGSPAGMSVQPKNWL 479
Db 421 PFHSSYAHQSLSLDRMLNPLIDQYLYLVRTQT-TGTGGTQTALFSAQAGPSSMANQARNWV 479
Qy 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db 480 PGPCYRQORVSTTTNQNNNSNFAWTGAAKFKLNGRDSLMPNPGVAMASHKODEDRFFPSSG 539
Qy 540 VMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAM 599
Db 540 VLIFGKQAGNDGVDYSQVLITDEEEIKATNPVATEEYGAVAINNQAANTQAOTGLVHNQ 599
Qy 600 GALPGMWQDRDQVYLGQPIWAKIPHTDGHFHPSPMLMGFGFLGNPPQILIKNTVPANPP 659
Db 600 GVIPGMVWQNRDQVYLGQPIWAKIPHTDGNFHPSPMLMGFGFLGNPPQILIKNTVPADPP 659
Qy 660 AEFSATKPFASITQYSTQGVSVSVEIEWELQENSKRWNPVQYTSNYAKSANVDFTVDNNG 719
Db 660 LTFNQAKLNSFITQYSTQGVSVSVEIEWELQENSKRWNPVQYTSNYAKSTNVDFAVNTEG 719
Qy 720 LYTEPRPIGTRYLTRPL 736
Db 720 VYSEPRPIGTRYLTRNL 736

RESULT 8
US-10-423-704A-7
; Sequence 7, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630


```

; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 9
US-10-423-704A-7

Query Match      86.9%; Score 3467; DB 12; Length 736;
Best Local Similarity 85.5%; Pred. No. 4.9e-293;
Matches 630; Conservative 44; Mismatches 61; Indels 2; Gaps 2;

Qy      1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60
Db      1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60

Qy      61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120
Db      61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120

Qy      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQPAKKRLNFGQTDSE 180
Db      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQPAKKRLNFGQTDSE 180

Qy      181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db      181 SVPDPQPLGEPPEAPSGLGPNTMASGGGAPMADNNNEGADGVGNSSGNWHCDSTWLGDRVI 240

Qy      241 TTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDENRFHCHFSPRDWQR 299
Db      241 TTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDENRFHCHFSPRDWQR 300

Qy      300 LINNWGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAH 359
Db      301 LINNWGFRPKRLNFKLFNIOQKEVTNEGTKTIANNLSTVQVFTDSEYQLPYVLGSAH 360

Qy      360 QGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEV 419
Db      361 QGCLPPFPADVFMVFPQYGYLTLNNGSQALGRSSFYCLEYFPPSQMLRTGNFTFSYTFEDV 420

Qy      420 PFHSSYAHSQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWL 479
Db      421 PFHSSYAHSQSLDRLMNPLIDQYLYLVRQT-TGTGGTQTTLAFSQAGPSSMANQARNWV 479

Qy      480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db      480 PGPCYRQORVSTTTNQNNSNFAWTGAAKFKLNGRDSLMPNPGVAMASHKODEDRFFPSSG 539

Qy      540 VMIFGESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAM 599
Db      540 VLIFGKQAGNDGVDYSQVLITDEEEIKATNPVATEEYGAVAINNQAAANTQAOTGLVHNQ 599

Qy      600 GALPGMWQDRDQVYLQGPVIAKIPHTDGHFHPSPMLMGGFGLKNPPPPQILIKNTPPVPANPP 659
Db      600 GVIPGMVQNRDQVYLQGPVIAKIPHTDGNFHPSPMLMGGFGLKHPPPPQILIKNTPPVPADPP 659

Qy      660 AEFSATKPFASFITQYSTGQVSVEIEWELQKNSKRWNPEVQYTSNYAKSANVDFTVDNNG 719
Db      660 LTFNQAKLNSFITQYSTGQVSVEIEWELQKNSKRWNPEIQYTSNYKSTNVDFAVNTEG 719

Qy      720 LYTEPRPIGTRYLTRPL 736
Db      720 VYSEPRPIGTRYLTRNL 736

RESULT 9
US-10-291-583-96
; Sequence 96, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
```

```

; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifie
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 43.21
US-10-291-583-96

Query Match      86.8%; Score 3464; DB 12; Length 736;
Best Local Similarity 85.3%; Pred. No. 8.9e-293;
Matches 629; Conservative 46; Mismatches 60; Indels 2; Gaps 2;

Qy      1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60
Db      1 MAADGYLPDWLEDNLSEGIREWMDLKPAPKPKANQQKDDGRGLVLPGYKYLGPFGNGLD 60

Qy      61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120
Db      61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGNGLGRAVFQ 120

Qy      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQPAKKRLNFGQTDSE 180
Db      121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQPAKKRLNFGQTDSE 180

Qy      181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db      181 SVPDPQPLGEPPEAPSGLGPNTMASGGGAPMADNNNEGADGVGNSSGNWHCDSTWLGDRVI 240

Qy      241 TTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDENRFHCHFSPRDWQR 299
Db      241 TTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDENRFHCHFSPRDWQR 300

Qy      300 LINNWGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAH 359
Db      301 LINNWGFRPKRLNFKLFNIOQKEVTNEGTKTIANNLSTVTRVFTDSEYQLPYVLGSAH 360

Qy      360 QGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEV 419
Db      361 QGCLPPFPADVFMVFPQYGYLTLNNGSQALGRSSFYCLEYFPPSQMLRTGNFTFSYTFEDV 420

Qy      420 PFHSSYAHSQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKOLLFSRGSAGMSVQPKNWL 479
Db      421 PFHSSYAHSQSLDRLMNPLIDQYLYLVRQT-TGTGGTQTTLAFSQAGPSSMANQARNWV 479

Qy      480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db      480 PGPCYRQORVSTTTNQSNNNSNFAWTGAAKFKLNGRDSLMPNPGVAMASHKDDDRFFPSSG 539

Qy      540 VMIFGESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAM 599
Db      540 VLIFGKQAGNDGVDYSQVLITDEEEIKATNPVATEEYGAVAINNQAAANTQAOTGLVHNQ 599

Qy      600 GALPGMWQDRDQVYLQGPVIAKIPHTDGHFHPSPMLMGGFGLKNPPPPQILIKNTPPVPANPP 659
Db      600 GVIPGMVQNRDQVYLQGPVIAKIPHTDGNFHPSPMLMGGFGLKHPPPPQILIKNTPPVPADPP 659

Qy      660 AEFSATKPFASFITQYSTGQVSVEIEWELQKNSKRWNPEVQYTSNYAKSANVDFTVDNNG 719
Db      660 LTFNQAKLNSFITQYSTGQVSVEIEWELQKNSKRWNPEIQYTSNYKSTNVDFAVNTEG 719
```

Qy	720	LYTEPRPIGTRYLTRPL	736
		: : : : : : : :	
Dd	720	VYSEPRPIGTRYLTRNL	736

RESULT 10

```

US-10-291-583-98
; Sequence 98, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 43.23
US-10-291-583-98

```

Query Match	86.7%;	Score 3460;	DB 12;	Length 736;
Best Local Similarity	85.3%;	Pred. NO. 2e-292;		
Matches 629;	Conservative 44;	Mismatches 62;	Indels 2;	Gaps 2;

QY	1	MAADGYLPDWLEDNLS	EGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLEDNLS	EGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD	60
QY	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADA	AEQERLQEDTSGGNLGRAVFQ	120
Db	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADA	AEQERLQEDTSGGNLGRAVFQ	120
QY	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSGIGKTGQQPAKKRLNFGQTGDSE	180	
Db	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSGIGKTGQQPAKKRLNFGQTGDSE	180	
QY	181	SVPDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNHCHDSTWLGD	RVI 240	
Db	181	SVPDPQPLGEPPAPSGLGPNTMASGGGAPMADNNEGADGVGNSSGNHCHDSTWLGD	RVI 240	
QY	241	TTSTRTWALPTYNNHLYKQISSASTGAS - NDNHYFGYSTPWGYFDFNRFHCHFS	PRDWQR 299	
Db	241	TTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDFNRFHCHFS	PRDWQR 300	
QY	300	LINNNWGFRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTVQVFS	DSYQLPYVLGSAH 359	
Db	301	LINNNWGFRPKRLNFKLFNIQVKEVTNEGTKTIANNLSTVQVFTD	LEYQLPYVLGSAH 360	
QY	360	QGCLPPFPADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGN	TFSYTFEEV 419	
Db	361	QGCLPPFPADVFMPQYGYLTLNNGSQALGRSSFYCLEYFPSQMPRTGN	TFSYTFEDV 420	
QY	420	PFHSSYAHSQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKOLLFSRGS	PAGMSVPQKNWL 479	
Db	421	PFHSSYAHSQSLDRLMNPLIDQYLYLVRQT - TGTGGTQT	LAFSQAGPSSMANQARNWV 479	
QY	480	PGPCYRQQRVSKTKTDNNNSNFTWTGASKYKYNLNGRESIIINPGTAMASH	KDDEDDKFFPMSG 539	

D b	480	PGPCYRQORVSTTTNQNNNSNFAWTGAAKFKLNGRDSLMPNPGVAMASHKDDDDDRFFPSSG	539
Q y	540	VMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPA ¹ TGDVHAM	599
D b	540	VLI ² FGKQGAGNDGVDYSQVLIITDEEEIKATNPVATEEYGA ³ VAINNQAANTQAQTGLVHNQ	599
Q y	600	GALPGMVQNRDRVYLQGP ⁴ PIWAKIPH ⁵ TDGHFHPSP ⁶ LMGGFGLKNP ⁷ PPQ ⁸ ILIK ⁹ NTVPVANPP	659
D b	600	GVIPGMVWQNRDRVYLQGP ⁴ PIWAKIPH ⁵ TDGNFHPSP ⁶ LMGGFGLKH ⁷ PPQ ⁸ ILIK ⁹ NTVPADPP	659
Q y	660	AEFSATKFASFITQYSTGQVSVEIEWELQKENS ¹ KRWNP ² EVQYTSNYAKS ³ ANV ⁴ FTVDNNG	719
D b	660	LTFNQA ⁵ KLNSFITQYSTGQVSVEIEWELQKENS ⁶ KRWNP ⁷ EIQYTSNY ⁸ YKSTNV ⁹ DFAVNTEG	719
Q y	720	LYTEPRPIGTRYLTRPL	736
D b	720	VYSEPRPIGTRYLTRNL	736

RESULT 11

```

US-10-291-583-99
; Sequence 99, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvirra, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 736
; TYPE: prt
; ORGANISM: capsid protein of AAV serotype, clone 43.20
US-10-291-583-99

```

Query Match	86.7%;	Score 3457;	DB 12;	Length 736;
Best Local Similarity	85.3%;	Pred. No. 3.6e-292;		
Matches 629;	Conservative	44;	Mismatches 62;	Indels 2;
				Gaps 2;

QY	1	MAADGYLPDWLEDNLSEGIREWWDLKPAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLEDNLSEGIREWWDLKPAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD	60
QY	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOFERLQEDTSFGGNLGRAVFQ	120
Db	61	KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOFERLQEDTSFGGNLGRAVFQ	120
QY	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGTKTQQQPAKKRLNFGOTGDSE	180
Db	121	AKKRVLEPLGLVEEGAKTAPGKKRLVEQSPQEPDSSSGIGTKTQQQPAKKRLNFGOTGDSE	180
QY	181	SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGD RVI	240
Db	181	SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGD RVI	240
QY	241	TTSTRTRTWPALPTYNHLYKQISSASTGAS - NDNHYFGYSTPWGYDFDNRFCHFPSPRDWQR	299
Db	241	TTSTRTRTWPALPTYNHLYKQISSASTGAS - NDNHYFGYSTPWGYDFDNRFCHFPSPRDWQR	300
QY	300	LINNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTTIANNLSTSTVQVFSDESSEYQLPYVLGSAH	359

Db 301 LINNNWGRPKRLNFKLFNIQVKEVTNNEGKTIANNLTSTVQVFTDSEYQLPYVLGSAH 360
Qy 360 OGCLPPFPADVFMIPOYGYLTNLNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEEV 419
Db 361 OGCLPPFPADVFTVPOYGYLTNLNGSQALGRSSFYCLEYFSPQMLRTGNFTFSYTFEDV 420
Qy 420 PFHSSYAHQSOLDRLMNPLIDQYLYLNRNQSGSAQNKDLLFSRGSFAGMSVQPKNWL 479
Db 421 PFHSSYAHQSOLDRLMNPLIDQYLYLVRQT-TGTGTTQLAFSQAGPSSMANQARNWV 479
Qy 480 PGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG 539
Db 480 PGPCYRQORVSTTTNQNNNSNFAWTGAAKFKLNGRDSLMPGVAMASHKDDDDRRFFPSSG 539
Qy 540 VMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAM 599
Db 540 VLIFGKQAGNDGVDYSQVLTITDEEEIKATNPVATEEYGAVAINNQAANTQAQTGLVHNQ 599
Qy 600 GALPGMWQDRDYYLQGPWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTPVPANPP 659
Db 600 GVIPGMVWQNRDYYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPPPPQILIKNTPVPADPP 659
Qy 660 AEFSATKFASFITQYSTGVSVIEWELQKNSKRWNPVEVQYTSNYAKSANVDFTVDNNG 719
Db 660 LTFNQAKLNSFITQYSTGVSVIEWELQKNSKRWNPVEIQYTSNYKSTNVDFAVNTEG 719
Qy 720 LYTEPRPIGTRYLTRPL 736
Db 720 VYSEPRPIGTRYLTRNL 736
RESULT 12
US-10-291-583-85
; Sequence 85, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilton, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 738
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 42.8
US-10-291-583-85
Query Match 86.6%; Score 3453; DB 12; Length 738;
Best Local Similarity 85.1%; Pred. No. 8.1e-292;
Matches 628; Conservative 45; Mismatches 63; Indels 2; Gaps 2;
Qy 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPYKYLGPENGLD 60
Db 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPYKYLGPENGLD 60
Qy 61 KGEVNAADAAAEHDKAYDQQLKAGDNPYLRYNHADAERFQERLQEDTSFGGNLGRAVFQ 120
Db 61 KGEVNAADAAAEHDKAYDQQLKAGDNPYLRYNHADAERFQERLQEDTSFGGNLGRAVFQ 120

Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQ-EPDSSSGIGKTGQOPAKKRLNFGQTGDS 179
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQSPDSSTGIGKTGQOPAKKRLNFGQTGDS 180
Qy 180 ESVPDPQPLGEPPTAPAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDV 239
Db 181 ESVPDPQPIGEPGAGPSGLSGTMAAGGAPMADNNEGADGVSSSGNWHCDSTWLGDV 240
Qy 240 ITTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYPDFNRFHCHFSRPRDW 298
Db 241 ITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYPDFNRFHCHFSRPRDW 300
Qy 299 RLINNNWGFPRKRLNFKLFNIQVKEVTNDGVTTIANNLTSTVQVFSDESQYQLPYVLGSA 358
Db 301 RLINNNWGFPRKRLNFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLGSA 360
Qy 359 HQGCLPPFPADVFMIPOYGYLTNLNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYTFEE 418
Db 361 HQGCLPPFPADVFMIPOYGYLTNLNGSQAVGRSSFYCLEYFSPQMLRTGNFTFSYQFED 420
Qy 419 VPFHSSYAHQSOLDRLMNPLIDQYLYLNRNQSGSAQNKDLLFSRGSFAGMSVQPKNWL 478
Db 421 VPFHSSYAHQSOLDRLMNPLIDQYLYLNRNQSGTAGTQQLLFSQAGPNNMSAQAKNW 480
Qy 479 LPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMS 538
Db 481 LPGPCYRQORVSTTLSONNNNSFAWTGATKYHLNGRDSLVPGVAMATHKDDDEERFFPSS 540
Qy 539 GVMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHA 598
Db 541 GVLMEFGKQAGKNDVYSSVMLTSEEEIKTNPVATEQYGVVADNLQQQNAAPIVGAVNS 600
Qy 599 MGALPGMWQDRDYYLQGPWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTPVPANP 658
Db 601 QGALPGMWQNRDYYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPPPPQILIKNTPVPADP 660
Qy 659 PAEFSATKFASFITQYSTGVSVIEWELQKNSKRWNPVEVQYTSNYAKSANVDFTVDN 718
Db 661 PTTFSQAKLASFITQYSTGVSVIEWELQKNSKRWNPVEIQYTSNYKSTNVDFAVNTE 720
Qy 719 GLYTEPRPIGTRYLTRPL 736
Db 721 GTYSEPRPIGTRYLTRNL 738
RESULT 13
US-10-291-583-79
; Sequence 79, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilton, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 738
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 44.1
US-10-291-583-79


```
Query Match      86.4%; Score 3448; DB 12; Length 738;
Best Local Similarity 85.0%; Pred. No. 2.2e-291;
Matches 627; Conservative 46; Mismatches 63; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
   |||
Db 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEQERLQEDTSFGNGLGRAVFQ 120
   |||
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEQERLQEDTSFGNGLGRAVFQ 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGQPAKKRLNFGQTGDS 179
   |||
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQSPQRSPTSSTGIGKGQPAKKRLNFGQTGDS 180

QY 180 ESVDPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDV 239
   |||
Db 181 ESVDPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDV 240

QY 240 ITTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDENRFHCHFSPRDWQ 298
   |||
Db 241 ITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDENRFHCHFSPRDWQ 300

QY 299 RLINNNGWFRPKRLNFKLFNIQVKEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSA 358
   |||
Db 301 RLINNNGWFRPKRLNFKLFNIQVKEVTQNEGKTIANNLSTIQVFTDSEYQLPYVLGSA 360

QY 359 HQGCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTTEE 418
   |||
Db 361 HQGCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYQFED 420

QY 419 VPFHSSYAHQSLSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKNW 478
   |||
Db 421 VPFHSSYAHQSLSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSQAGPNNMSAQAKNW 480

QY 479 LPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMS 538
   |||
Db 481 LPGPCYRQQRVSTTLSQNNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDDEERFFPSS 540

QY 539 GVMIFGKESAGASNTALDNNVITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA 598
   |||
Db 541 GVLMPGKQAGKDNVDYSSVMLTSEEEIKTNPVATEQYGVVADNLQQQNAAPIVGAVNS 600

QY 599 MGALPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPIILIKNTPV PANP 658
   |||
Db 601 QGALPGMWQNRDVLQGPPIWAKIPHTDGNFHPSPLMGGFGLKHPPPIILIKNTPV PADP 660

QY 659 PAEFSATKFAFITQYSTQGVSVIEIWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNN 718
   |||
Db 661 PTTFSQAKLASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYKSTNVDFAVNTD 720

QY 719 GLYTEPRPIGTRYLTRPL 736
   |||
Db 721 GTYSEPRPIGTRYLTRNL 738

RESULT 14
US-10-291-583-81
; Sequence 81, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1 Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
```

```
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 738
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 44.2
US-10-291-583-81

Query Match      86.4%; Score 3448; DB 12; Length 738;
Best Local Similarity 85.0%; Pred. No. 2.2e-291;
Matches 627; Conservative 46; Mismatches 63; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
   |||
Db 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEQERLQEDTSFGNGLGRAVFQ 120
   |||
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEQERLQEDTSFGNGLGRAVFQ 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGQPAKKRLNFGQTGDS 179
   |||
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQSPQRSPTSSTGIGKGQPAKKRLNFGQTGDS 180

QY 180 ESVDPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDV 239
   |||
Db 181 ESVDPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDV 240

QY 240 ITTSTRTWALPTYNNHLYKQISSASTGAS-NDNHYFGYSTPWGYFDENRFHCHFSPRDWQ 298
   |||
Db 241 ITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDENRFHCHFSPRDWQ 300

QY 299 RLINNNGWFRPKRLNFKLFNIQVKEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSA 358
   |||
Db 301 RLINNNGWFRPKRLNFKLFNIQVKEVTQNEGKTIANNLSTIQVFTDSEYQLPYVLGSA 360

QY 359 HQGCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTTEE 418
   |||
Db 361 HQGCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYQFED 420

QY 419 VPFHSSYAHQSLSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKNW 478
   |||
Db 421 VPFHSSYAHQSLSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSQAGPNNMSAQAKNW 480

QY 479 LPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMS 538
   |||
Db 481 LPGPCYRQQRVSTTLSQNNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDDEERFFPSS 540

QY 539 GVMIFGKESAGASNTALDNNVITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA 598
   |||
Db 541 GVLMPGKQAGKDNVDYSSVMLTSEEEIKTNPVATEQYGVVADNLQQQNAAPIVGAVNS 600

QY 599 MGALPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKNPPPIILIKNTPV PANP 658
   |||
Db 601 QGALPGMWQNRDVLQGPPIWAKIPHTDGNFHPSPLMGGFGLKHPPPIILIKNTPV PADP 660

QY 659 PAEFSATKFAFITQYSTQGVSVIEIWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNN 718
   |||
Db 661 PTTFSQAKLASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYKSTNVDFAVNTD 720

QY 719 GLYTEPRPIGTRYLTRPL 736
   |||
Db 721 GTYSEPRPIGTRYLTRNL 738

RESULT 15
US-10-291-583-93
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 15:59:46 ; Search time 44.8925 Seconds
(without alignments)
2602.280 Million cell updates/sec

Title: US-09-807-802A-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSGIR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :					A_Geneseq_19Jun03.*				
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*					1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*				
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*					2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*				
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*					3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*				
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*					4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*				
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*					5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*				
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*					6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*				
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*					7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*				
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*					8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*				
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*					9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*				
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*					10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*				
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*					11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*				
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*					12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*				
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*					13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*				
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*					14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*				
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*					15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*				
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*					16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*				
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*					17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*				
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*					18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*				
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*					19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*				
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*					20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*				
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*					21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*				
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*					22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*				
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*					23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*				
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*					24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*				
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
					SUMMARIES				
Result No.	Score	Query Match	Length	ID	Description				
-----					-----				
1	3989	100.0	736	21	AAV71167	Adeno-associated v			
2	3963	99.3	736	22	AAV59847	AAV6 capsid protei			
3	3511	88.0	736	22	AAV59846	AAV3B capsid prote			
4	3494	87.6	736	22	AAV59845	AAV3A capsid prote			
5	3402.5	85.3	735	22	AAV65792	Adeno-associated v			
6	3402.5	85.3	735	22	AAV51508	Adeno-associated v			
7	3402.5	85.3	735	22	AAV59844	AAV2 capsid protei			
8	3402.5	85.3	735	23	AAU98974	Adeno-associated v			
9	3376	84.6	734	22	AAV50326	Adeno-associated v			

10	3251	81.5	599	21	AAV71168	Adeno-associated v
11	2906	72.9	534	21	AAV71169	Adeno-associated v
12	2759.5	69.2	598	22	AAG65793	Adeno-associated v
13	2759.5	69.2	598	22	AAM51509	Adeno-associated v
14	2759.5	69.2	598	23	AAU98975	Adeno-associated v
15	2486.5	62.3	734	19	AAW46308	AAV4 VP1 capsid pr
16	2486.5	62.3	734	24	ABG73935	Adeno-associated v
17	2481.5	62.2	533	22	AAG65794	Adeno-associated v
18	2481.5	62.2	533	23	AAU98976	Adeno-associated v
19	2477.5	62.1	533	22	AAM51510	Adeno-associated v
20	2220	55.7	724	21	AAV58160	Adeno-associated v
21	2220	55.7	724	23	AAU11405	Adeno-associated v
22	2195.5	55.0	736	18	AAW19000	Duck parvovirus ca
23	2168.5	54.4	732	16	AAR85385	Barbary duck parvo
24	1830.5	45.9	598	19	AAW46313	AAV4 VP2 coat prot
25	1830.5	45.9	598	24	ABG73940	Adeno-associated v
26	1700.5	42.6	588	21	AAV58161	Adeno-associated v
27	1700.5	42.6	588	23	AAU11406	Adeno-associated v
28	1690.5	42.4	544	19	AAW46314	AAV4 VP3 coat prot
29	1690.5	42.4	544	24	ABG73941	Adeno-associated v
30	1676	42.0	534	16	AAR85386	Barbary duck parvo
31	1665	41.7	532	21	AAV58162	Adeno-associated v
32	1665	41.7	532	23	AAU11407	Adeno-associated v
33	1410	35.3	756	21	AAV71231	Capsid protein enc
34	921	23.1	202	23	AAU83481	Adeno-associated v
35	606.5	15.2	781	20	AAV23227	Erythrovirus V9 VP
36	605.5	15.2	781	16	AAW08986	Human parvovirus V
37	599.5	15.0	781	24	ABP57263	Human parvovirus B
38	590.5	14.8	781	24	ABP57266	Human parvovirus B
39	490	12.3	554	16	AAW08987	Human parvovirus V
40	490	12.3	554	20	AAV23230	Erythrovirus V9 VP
41	486	12.2	554	24	ABP57264	Human parvovirus B
42	486	12.2	554	24	ABP57267	Human parvovirus B
43	476.5	11.9	543	12	AAP13405	Parvo virus B19 VP
44	425.5	10.7	472	5	AAP40675	Sequence encoded b
45	417.5	10.5	620	22	AAM50111	Feline parvovirus

ALIGNMENTS

RESULT 1				
AAV71167				
ID	AAV71167 standard; Protein; 736 AA.			
XX				
AC	AAV71167;			
XX				
DT	08-SEP-2000 (first entry)			
XX				
DE	Adeno-associated virus serotype 1 capsid protein VP1.			
XX				
KW	Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; vaccine; transgene; VP1.			
KW				
XX				
OS	Adeno associated virus serotype 1.			
XX				
PN	WO200028061-A2.			
XX				
PD	18-MAY-2000.			
XX				
PF	02-NOV-1999; 99WO-US25694.			
XX				
PR	05-NOV-1998; 98US-0107114.			
XX				
PA	(UYPE-) UNIV PENNSYLVANIA.			
XX				
PI	Wilson JM, Xiao W;			
XX				
DR	WPI; 2000-376571/32.			
DR	N-PSDB; AAD00772, AAD00777.			
XX				
PT	Novel adeno-associated virus serotype 1 polynucleotide useful for			

PT preparation of medicament for delivery of a transgene to a host -
XX
PS Claim 7; Page 87-90; 108pp; English.
XX

CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transfecting host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 cap protein VP1 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX

SQ Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 21; Length 736;
Best Local Similarity 100.0%; Pred. No. 1e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD 60
QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEFQERLQEDTSFGNGLGRAVFQ 120
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEFQERLQEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
QY 181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRL 300
QY 301 INNWGFRPKRLNFKLFNIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
Db 301 INNWGFRPKRLNFKLFNIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
Db 361 GCLPPFPADVFMIPQYGYLTNLNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVP 420
QY 421 FHSSYAHQSQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLP 480
Db 421 FHSSYAHQSQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEKFFPMSGV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDEKFFPMSGV 540
QY 541 MIFGKESAGASNTALDNVMTDDEEIKATNPVATERFGTVAVNFFQSSSTDPAATGDVHAMG 600
Db 541 MIFGKESAGASNTALDNVMTDDEEIKATNPVATERFGTVAVNFFQSSSTDPAATGDVHAMG 600
QY 601 ALPGMWQDRDQVYLGPIWAKIPHTDGHFHPSPLMGGFGGLKNPPPPQILIKNTPVPANPPA 660
Db 601 ALPGMWQDRDQVYLGPIWAKIPHTDGHFHPSPLMGGFGGLKNPPPPQILIKNTPVPANPPA 660
QY 661 EFSATKFPASFITQYSTGQVSVEIEWELQENSKRWNPVEQVYTSNYAKSANVDFTVDNNGL 720
Db 661 EFSATKFPASFITQYSTGQVSVEIEWELQENSKRWNPVEQVYTSNYAKSANVDFTVDNNGL 720

QY 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736
RESULT 2
AAB59847
ID AAB59847 standard; Protein; 736 AA.
XX
AC AAB59847;
XX
DT 28-MAR-2001 (first entry)
XX
DE AAV6 capsid protein VP1.
XX
KW AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.
XX
OS Adeno associated virus.
XX
PN US6156303-A.
XX
PD 05-DEC-2000.
XX
PF 11-JUN-1997; 97US-0873168.
XX
PR 11-JUN-1997; 97US-0873168.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Russell DW, Rutledge EA;
XX
DR WPI; 2001-060164/07.
XX
PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -
XX
PS Disclosure; Fig 2; 50pp; English.
XX

CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV6). AAV6
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX
SQ Sequence 736 AA;

Query Match 99.3%; Score 3963; DB 22; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.2e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDNLSGIREWMDLKPAGPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD 60
QY 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEFQERLQEDTSFGNGLGRAVFQ 120
Db 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAFAEFQERLQEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSE 180
QY 181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240

Qy	241	TTSTR	TWALPT	YNNHLYKQISSASTGASNDNHFGYSTPWCYFD	FNRFCHFS	PRDWQRL	300		
Db	241	TTSTR	TWALPT	YNNHLYKQISSASTGASNDNHFGYSTPWCYFD	FNRFCHFS	PRDWQRL	300		
Qy	301	INNNG	FRPKRL	NFKLFNIOQKEVTTNDGVT	TTIANNLSTVQV	SDSEYQLPYVLGSAHQ	360		
Db	301	INNNG	FRPKRL	NFKLFNIOQKEVTTNDGVT	TTIANNLSTVQV	SDSEYQLPYVLGSAHQ	360		
Qy	361	GCLPP	FPADV	FMIPQYGYLT	LNNGSQAVGRSSFC	LEYFPPSOMLRTGN	NFTFSYTFEEVP	420	
Db	361	GCLPP	FPADV	FMIPQYGYLT	LNNGSQAVGRSSFC	LEYFPPSOMLRTGN	NFTFSYTFEDVP	420	
Qy	421	FHSSY	AHSQSL	DRLMNP	LIDQYLYLNRTNQSG	SAQNKDLLFSRGS	PAGMSVQPKNWLP	480	
Db	421	FHSSY	AHSQSL	DRLMNP	LIDQYLYLNRTNQSG	SAQNKDLLFSRGS	PAGMSVQPKNWLP	480	
Qy	481	GPCYR	QQRVSK	TKTDN	NNSNFTWTGASKYN	LNGRESIINPGT	AMASHKDDKFFPMSGV	540	
Db	481	GPCYR	QQRVSK	TKTDN	NNSNFTWTGASKYN	LNGRESIINPGT	AMASHKDDKFFPMSGV	540	
Qy	541	MIFGK	ESAGAS	NTALDN	VMTDEEEIKATNP	VATERFGTVAVNF	QSSSTDPA	TGDVHAMG	600
Db	541	MIFGK	ESAGAS	NTALDN	VMTDEEEIKATNP	VATERFGTVAVNF	QSSSTDPA	TGDVHVMG	600
Qy	601	ALPGM	VWQDR	DVYLQGP	PIWAKIPHTDGHF	HPSPLMGGFGLKN	PPQILIKNT	PVPANPPA	660
Db	601	ALPGM	VWQDR	DVYLQGP	PIWAKIPHTDGHF	HPSPLMGGFGLKN	PPQILIKNT	PVPANPPA	660
Qy	661	EPSAT	KFASFI	TQYSTG	VSVIEWELQKENS	KRWNP	EVQYTSNYAKS	ANVDFTVDNNGL	720
Db	661	EPSAT	KFASFI	TQYSTG	VSVIEWELQKENS	KRWNP	EVQYTSNYAKS	ANVDFTVDNNGL	720
Qy	721	YTEPR	PIGTRY	LTRPL	736				
Db	721	YTEPR	PIGTRY	LTRPL	736				
RESULT 3									
AAB59846									
ID	AAB59846 standard; Protein; 736 AA.								
XX	AAB59846;								
AC	28-MAR-2001 (first entry)								
XX	AAB59846;								
DT	28-MAR-2001 (first entry)								
XX	AAB59846;								
DE	28-MAR-2001 (first entry)								
XX	AAB59846;								
KW	AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;								
KW	atherosclerosis; sickle cell anaemia; thalassaemia;								
KW	blood clotting disorder; diabetes; capsid protein VP1.								
XX	Adeno associated virus.								
OS	US6156303-A.								
XX	05-DEC-2000.								
PD	11-JUN-1997; 97US-0873168.								
XX	11-JUN-1997; 97US-0873168.								
PR	(UNIW) UNIV WASHINGTON.								
XX	Russell DW, Rutledge EA;								
PI	WPI; 2001-060164/07.								
XX	Adeno-associated virus serotype 6 and viral vector derived from it for								
DR	gene therapy of cystic fibrosis, cancer, acquired immunodeficiency								
XX	syndrome, sickle cell anemia, thalassemia and diabetes								
PS	Disclosure; Fig 2; 50pp; English.								
XX									

CC	The present invention relates to adeno-associated virus serotypes. The					
CC	present sequence is capsid protein VP1 of one such serotype (AAV3B).					
CC	AAV3B can be used to construct AAV viral vectors for use in gene therapy					
CC	for a range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,					
CC	sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.					
CC	The AAV viral vectors have increased transduction efficiency of a					
CC	particular host cell as the AAV virion containing the AAV vector genome					
CC	can be modified to express a capsid protein of an AAV serotype that					
CC	transduces the selected host cell.					
XX						
SQ	Sequence 736 AA;					
	Query Match	88.0%;	Score 3511;	DB 22;	Length 736;	
	Best Local Similarity	86.8%;	Pred. No. 2.8e-273;			
	Matches 640;	Conservative 38;	Mismatches 57;	Indels 2;	Gaps 2;	
QY	1	MAADGYLPDWLEDNL	SEGIREWDLKPGAPKPKANQKQDDGRGLV	LPGYKYLGPFGNGLD	60	
Db	1	MAADGYLPDWLEDNL	SEGIREWDLKPGVPKPKANQKHQDNRRLGLV	LPGYKYLGPFGNGLD	600	
QY	61	KGEPVNAADAAALEHDKAYDQQLKAGDNP	LYRYNHADA	EFQERLQEDTSFGGNLGRAVFQ	120	
Db	61	KGEPVNEADAAALEHDKAYDQQLKAGDNP	LYRYNHADA	EFQERLQEDTSFGGNLGRAVFQ	120	
QY	121	AKKRVLEPLGLVEE	GAKTAPGKKRPVEQSPQEPDSSSGIGKTQ	QPAKRLNFGQTGDSE	180	
Db	121	AKKRILEPLGLVEE	AAKTAPGKKRPVDQSPQEPDSSSGVKGSKQ	PARKRLNFGQTGDSE	180	
QY	181	SVDPDPQLGEP	PATPAAVGPTTMA	SGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240	
Db	181	SVDPDPQLGEP	PAPTSLGSNTMA	SGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRVI	240	
QY	241	TTSTRTWALPTYNNHLYKQISSASTGAS	NDNHFGYSTPWGYFDFNRFCHFS	PRDWQRL	300	
Db	241	TTSTRTWALPTYNNHLYKQISSQS-GAS	NDNHFGYSTPWGYFDFNRFCHFS	PRDWQRL	299	
QY	301	INNNGFRPKRLNFKLFNIOQKEVTNDGVT	TIANNLTSTVQV	SDSEYQLPYVLGSAHQ	360	
Db	300	INNNGFRPKLSFKLFNIOQKEVTQNDGTT	TIANNLTSTVQV	TDSEYQLPYVLGSAHQ	359	
QY	361	GCLPPFPADVFMIPQYGYLT	LNNGSQAVGRSSFCLEYFPPSOMLRTG	NNFTFSYTFEEVP	420	
Db	360	GCLPPFPADVFMVPQYGYLT	LNNGSQAVGRSSFCLEYFPPSOMLRTG	NNFQFSYTFEDVP	419	
QY	421	FHSSYAHSQS	LDRLMNP	LIDQYLYLNRTQ-NQSGSAQNKDLLFSR	GSPAGMSVQPKNWL 479	
Db	420	FHSSYAHSQS	LDRLMNP	LIDQYLYLNRTQGTTS	GTNQSRLLF	FSQAGPQMSLQARNWL 479
QY	480	PGPCYRQQRVSKTKTDN	NNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSG	539		
Db	480	PGPCYRQORLSKTANDN	NNSNFPWTAASKYHLNGRDSLVP	PGPAMASHKODEEKFPPMHG	539	
QY	540	VMI	FGKESAGASNTALDNVMTDEEEIKATNP	VATERFGTVAVNFQSSSTDPA	TGDVHAM 599	
Db	540	NLIFGKEGTTASNAELDNVMTDEEEIR	TNPVATEQYGTVAN	LLQSSNTAPTTRTVNDQ	599	
QY	600	GALPGMVWQDRDVYLOQPIWAKI	PHTDGHFHPSP	LMGGFGLKNPPQILIKNT	PVPANPP 659	
Db	600	GALPGMVWQDRDVYLOQPIWAKI	PHTDGHFHPSP	LMGGFGLKHPPPQIMIKNT	PVPANPP 659	
QY	660	AEFSATKFASFI	TOYSTGQVSVEIEWELQKNSKRWNPEVQY	TSNYAKSANVDFTVDNNG	719	
Db	660	TTFSPAKFASFI	TOYSTGQVSVEIEWELQKNSKRWNPEIQY	TSNYKSNVNDFTVDNTNG	719	
QY	720	LYTEPRPIGTRYL	TRPL	736		
Db	720	VYSEPRPIGTRYL	TRNL	736		
RESULT 4						
AAB59845						
ID	AAB59845 standard; Protein; 736 AA.					
XX						

AC AAB59845;
XX 28-MAR-2001 (first entry)
XX AAV3A capsid protein VP1.
DE AAV3A; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
XX atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.
XX Adeno associated virus.
OS US6156303-A.
XX 05-DEC-2000.
XX 11-JUN-1997; 97US-0873168.
XX 11-JUN-1997; 97US-0873168.
XX (UNIW) UNIV WASHINGTON.
XX Russell DW, Rutledge EA;
PI WPI; 2001-060164/07.
XX Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -
XX Disclosure; Fig 2; 50pp; English.
XX The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV3A).
CC AAV3A can be used to construct AAV viral vectors for use in gene therapy
CC for a range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX
SQ Sequence 736 AA;
Query Match 87.6%; Score 3494; DB 22; Length 736;
Best Local Similarity 86.4%; Pred. No. 6.6e-272;
Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;
Qy 1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLPGPENGLD 60
Db 1 MAADGYLPDWLEDNLSEGIREWDLKPGVPQPKANQQHQDNRRLVLPGYKYLPGPENGLD 60
Qy 61 KGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAERLQEDTSFGGNLGRAVFQ 120
Db 61 KGEPVNEADAAALEHDKAYDQQLKAGDNPYLKYNHADAERLQEDTSFGGNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKGTGQOPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEAAKTAPGKGAVDQSPQEPDSSSGVSGSKQPARKRLNFGQTGDSE 180
Qy 181 SVPDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI 240
Db 181 SVPDPQPLGEPPAAPTSLGSNTMASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRVI 240
Qy 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWGVDNFRFCHFSRPRDWQRL 300
Db 241 TTSTRTWALPTYNNHLYKQISSQS-GASNDNHFGYSTPWGVDNFRFCHFSRPRDWQRL 299
Qy 301 INNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQ 360
Db 300 INNNWGRPKKLSFKLFNIQVRGVTVQNDGTTIANNLTSTVQFTDSEYQLPYVLGSAHQ 359
Qy 361 GCLPPFPADVFMIPQYGLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVP 420

Db 360 GCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEDVP 419
Qy 421 FHSSYAHQSQSLDRLMNPLIDQYLYLNRQTQ-NQSGSAQNKDLLFSRGSAGMSVQPKNWL 479
Db 420 FHSSYAHQSQSLDRLMNPLIDQYLYLNRQTQTSGTTNQSRLLFSQAGPQSMQSLQARNWL 479
Qy 480 PGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMSG 539
Db 480 PGPCYRQQRLSKTANDNNNSNFPWTAASKYHLNGRDSLVPNPGPAMASHKDDDEKFFPMSG 539
Qy 540 VMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFSSTDPATGDVHAM 599
Db 540 NLIFGKEGTTASNAELDNVMITDEEEIRTNPVATEQYGTVANLQSSNTAPTGTVNHQ 599
Qy 600 GALPGMVWQDRDVYLGQPIWAKIPHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPP 659
Db 600 GALPGMVWQDRDVYLGQPIWAKIPHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPP 659
Qy 660 AEFSA TKFASFITQYSTGQSVSVEIEWELQKNSKRWNPVQYTSNYAKSANVDFTVDNNG 719
Db 660 TTFSPAKFASFITQYSTGQSVSVEIEWELQKNSKRWNPVQYTSNYAKSANVDFTVDNNG 719
Qy 720 LYTEPRPIGTRYLTRPL 736
Db 720 VYSEPRPIGTRYLTRNL 736
RESULT 5
AAG65792
ID AAG65792 standard; Protein; 735 AA.
XX AAG65792;
AC AAG65792;
XX 11-FEB-2002 (first entry)
DT Adeno-associated virus 2 (AAV-2) major coat protein VP1.
XX
DE Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytosstatic;
XX inverted terminal repeat; nootropic; neuroprotective; antianemic; ITR;
KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV;
KW major coat protein; AAV-2; VP1.
XX
XX Adeno-associated virus 2.
OS WO200168888-A2.
XX 20-SEP-2001.
XX 13-MAR-2001; 2001WO-US07927.
XX 14-MAR-2000; 2000US-189110P.
XX (NEUR-) NEUROLOGIX INC.
XX Xiao W, During MJ;
XX WPI; 2001-596912/67.
DR N-PSDB; AAI66974.
XX
PT Recombinant viral vector useful in improving gene therapy in a subject,
PT and for increasing efficiency of entry into a cell, comprises a
PT chimeric capsid having one non-native amino acid sequence and a desired
PT transgene -
XX Disclosure; Page 50; 53pp; English.
XX
CC The invention provides a recombinant viral vector (RVV) comprising a
CC chimeric capsid (I) having at least one non-native amino acid sequence,
CC derived from a capsid protein domain of parvovirus (II), a virus (III),
CC or their combination, and a transgene flanked 5' and 3' by inverted
CC terminal repeat (ITR) sequences, derived from (II), (III), or their
CC combination. The RVV is useful for improving gene therapy in a subject

with a disorder, and for increasing the efficiency of entry into a cell, which involves producing (I) encapsulating a viral vector, and contacting a cell with RVV having (I) such that (I) binds to an attachment site on the cell surface and permits the vector to enter the cell efficiently. A pharmaceutical composition comprising RVV with (I) containing a transgene sequence associated with a disease or a disorder such that expression of the transgene would result in amelioration of the disease or disorder such as inherited neurological and metabolic diseases e.g. lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polyneuropathy, Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic fibrosis, diabetes, diseases associated with hormone deficiencies, retinoblastoma and various types of neoplastic cells which include tumours especially central nervous system tumours, neoplasms, carcinomas, sarcomas, leukemias and lymphoma. The present sequence represents the adeno-associated virus 2 (AAV-2) major coat protein VP1. AAV-2 sequences are used in the construction of a chimeric vector.

Sequence	735 AA;
SQ	

Query Match 85.3%; Score 3402.5; DB 22; Length 735;
Best Local Similarity 83.3%; Pred. NO. 1.5e-264;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

Qy	1	MAADGYLPDWLEDNLS	EGIREWMDLKPGAPKPKANQQKQDDGRGLVLP	PGYKYLGPPFNGLD	60
Db	1	MAADGYLPDWLEDTL	SEGIQWKKLPGPPPKPAERHKDDSRGLVLP	PGYKYLGPPFNGLD	60
Qy	61	KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADA	AEFQERLQEDTSFGGNLGRAVFQ	120	
Db	61	KGEPVNEADAAALEHDKAYDRQLDSDGNPYLKYNHADA	AEFQERLXEDTSFGGNLGRAVFQ	120	
Qy	121	AKGRVLEPLGLVEEGAKTAPGKKRPVEQSPOEPDSSGIGK	TGQQPAKKRLNFGQTGDSE	180	
Db	121	AKGRVLEPLGLVEEPPVKTAPGKKRPVEHSPVEPDSSSGTG	KAGQQPARKRLNFGQTGDAD	180	
Qy	181	SVPDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGV	GNASGNWHCDSTWLGD RVI	240	
Db	181	SVPDPQPLGQPPAAPSGLTGNTMATGSGAPMADNNEGADGV	GNSSGNWHCDSTWMGD RVI	240	
Qy	241	TTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWG	YFDNRFHCHFSPRDWQRL	300	
Db	241	TTSTRTWALPTYNNHLYKQISSQS - GASNDNHYFGYSTPWG	YFDNRFHCHFSPRDWQRL	299	
Qy	301	INNNGFRPKRLNFKLENIQVKEVTNDGVTTIANNLSTVQV	FSDEYQLPYVLGSAHQ	360	
Db	300	INNNGFRPKRLNFKLENIQVKEVTQNDGTTTIANNLSTVQV	FTDSEYQLPYVLGSAHQ	359	
Qy	361	GCLPPFPADVFMIPOYGYLTNLNGSQAVGRSSFYCYLEY	FFSQMLRTGNNTFSYTFEEVP	420	
Db	360	GCLPPFPADVFMPVQGYLTNLNGSQAVGRSSFYCYLEY	FFSQMLRTGNNTFSYTFEDVP	419	
Qy	421	FHSSYAHSQSLDRLMNPLIDQYLYLNR	TQNQSGSAQNKKOLLFSRGS	PAGMSVQPKNWLP	480
Db	420	FHSSYAHSQSLDRLMNPLIDQYLYLSR	TNTPSGTTTQSFQSQAGASDIRDQSRNWLP	479	
Qy	481	GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESI	INPGTAMASHKDDDEKFFPMSGV	540	
Db	480	GPCYRQQRVSKTSADNNNSNFTWTGATKYHLNGRDSL	VNPGPAMASHKDDDEKFFPQSGV	539	
Qy	541	MIFGESAGASNTALDNVMTDDEEIKATNPVATERFGTV	AVNFQSSSTDPATGDVHAMG	600	
Db	540	LIFGKQSGSEKTNVDIEKVMITDEEIR	TNPVATEQYGSVSTNLQRGNRQAATADVNTQG	599	
Qy	601	ALPGMVWQDRDVYLOQPIWAKIPHTDGHFHPSP	MLMGFGFLKNPPPILIKNTVPANPPA	660	
Db	600	VLPGMVWQDRDVYLOQPIWAKIPHTDGHFHPSP	MLMGFGFLKHPPPILIKNTVPANPST	659	
Qy	661	EPSATKFPASFITQYSTGQSVSEIEWELQKENS	KRWNPVQYTSYNAKSANVDFTVDDNGL	720	
Db	660	TFSAAKFPASFITQYSTGQSVSEIEWELQKENS	KRWNPVQYTSYNAKSANVDFTVDTNGV	719	
Qy	721	YTEPRPIGTRYLTRPL	736		

```

db
720 YSEPRPIGTRYLTRNL 735
|:|||||

```

RESULT 6

AAM51508

ID	AAM51508 standard; protein; 735 AA.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AC AAM51508:

DT 02-JAN-2002 (first entry)

XXV
DE Adeno-associated virus VP1 capsid protein

Adeno-associated virus; AAV; VP1; capsid; virus-like particle; VP3. nuclear localisation signal; VP3.

OS Adeno associated virus.

PN JP2001169777-A.

26-JUN-2001

30-III.-1999. 99.1P-0249140

XX 30-III-1999. 987B-0349140

XX
PA
(HAND /) HANDA H[illegible]

PT New virus-like particles from VP3 capsid protein of adeno-associated
PT virus, comprise a peptide containing a nucleus-shifting signal
PT connected to its N-terminal -
XX

PS Disclosure; Page 10-13; 33pp; Japanese.

The present sequence is provided in a specification relating to a virus-like particle-forming protein, and to a peptide containing a nuclear-shifting signal at its N-terminus that can form virus-like particles by shifting to the nucleus of the animal cell in which it is expressed. The method is used for forming virus-like particles from the vp3 capsid protein of adeno-associated virus (AAV).

Sequence 735 AA;

Query Match	85.3%	Score 3402.5	DB 22	Length 735

Best Local Similarity 83.3%; Pred. No. 1.5e-264;

Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

QY 1 MAADGYLPDWLEDNLSEGIREWWDLKPGAEPKPKANOOKODDGRGLVLPGYKYLGPFFNGLD 60

db 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPPPKPAERHKDDSRGLVLPGYKYLGPFENGLD 60

61 KGEFVNAADAAALEHDKAYDOOLKAGDNPYLRYNHADAEEFOERLOEDTSEGGNI GRAVFO 120

61 KGEPVNEADAALEHDKAYNROI.DSGDNPI.KYNHADAEECEPI.KENTSEEGNI.GRAVEO 120

0v 121 AKKPVIFRPIGLVFECAKTA DGIKTPVFEVSDFEFDSSCTGKTGQDRAVYRI NEGOMCNSE 180

121 NKWIV FBI CI VEBBIVATA DCVKB DIBVUS DIBVBS SCCTCKV COODABVBYI NESCOTCDBB 180

181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 116

[illegible]

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 361 GCLPPFPADVFMIPQYGYLTILNNGSQAVGRSSFYCLEYFPQSMLRTGNNFTFSYTFEEVP 420
Db 360 GCLPPFPADVFMVPQYGYLTILNNGSQAVGRSSFYCLEYFPQSMLRTGNNFTFSYTFEDVP 419
QY 421 FHSSYAHQSQSLDRLMNPIDQYLYLNRQTQNSGAQNKDLLFSRGSPAGMSVQPKNWLP 480
Db 420 FHSSYAHQSQSLDRLMNPIDQYLYLNRQTQNSGAQNKDLLFSRGSPAGMSVQPKNWLP 479
QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGV 540
Db 480 GPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKDEKFFPMSGV 539
QY 541 MIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
Db 540 LIFGKQGSEKTNVDIEKVMITDEEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQG 599
QY 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKNPPQIILIKNTVPVPPANPPA 660
Db 600 VLPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKHPPPPQIILIKNTVPVPPANPST 659
QY 661 EFSATKFASFITQYSTGQVSVEIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNNGL 720
Db 660 TFSAAKFASFITQYSTGQVSVEIEWELQKNSKRWNPVEIQYTSNYKSNVVDFTVDTNVG 719
QY 721 YTEPRPIGTRYLTRPL 736
Db 720 YSEPRPIGTRYLTRNL 735

RESULT 7
AAB59844
ID AAB59844 standard; Protein; 735 AA.
XX
AC AAB59844;
XX

DT 28-MAR-2001 (first entry)
XX
DE AAV2 capsid protein VP1.
XX

KW AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.

OS Adeno associated virus.

XX US6156303-A.

PD 05-DEC-2000.

XX 11-JUN-1997; 97US-0873168.

PR 11-JUN-1997; 97US-0873168.

XX (UNIW) UNIV WASHINGTON.

PI Russell DW, Rutledge EA;

XX WPI; 2001-060164/07.

PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -

XX Claim 7; Fig 2; 50pp; English.

PS The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV2). AAV2
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome

CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX
SQ Sequence 735 AA;

Query Match 85.3%; Score 3402.5; DB 22; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.5e-264;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

QY 1 MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEPPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEEFQERLQEDTSFGGNLGRAVFP 120
Db 61 KGEPPVNEADAAALEHDKAYDRQLDSGDNPYLRKYNHADAEEFQERLKEDTSFGGNLGRAVFP 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKRPVEQSPQEPDSSSGIGTKTQQPAKKRLNFGQTGDSE 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLGEPPTAPAAVGTPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDVFI 240
Db 181 SVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDGDRVI 240

QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSFPRDWQL 300
Db 241 TTSTRTWALPTYNNHLYKQISSQS - GASNDNHFGYSTPWGYFDNRFCHFSFPRDWQL 299

QY 301 INNNWGFPRKRLNFKLFNIQVKEVTNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
Db 300 INNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQ 359

QY 361 GCLPPFPADVFMIPQYGYLTILNNGSQAVGRSSFYCLEYFPQSMLRTGNNFTFSYTFEEVP 420
Db 360 GCLPPFPADVFMVPQYGYLTILNNGSQAVGRSSFYCLEYFPQSMLRTGNNFTFSYTFEDVP 419

QY 421 FHSSYAHQSQSLDRLMNPIDQYLYLNRQTQNSGAQNKDLLFSRGSPAGMSVQPKNWLP 480
Db 420 FHSSYAHQSQSLDRLMNPIDQYLYLNRQTQNSGAQNKDLLFSRGSPAGMSVQPKNWLP 479

QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGV 540
Db 480 GPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKDEKFFPMSGV 539

QY 541 MIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
Db 540 LIFGKQGSEKTNVDIEKVMITDEEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQG 599

QY 601 ALPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKNPPQIILIKNTVPVPPANPPA 660
Db 600 VLPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKHPPPPQIILIKNTVPVPPANPST 659

QY 661 EFSATKFASFITQYSTGQVSVEIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNNGL 720
Db 660 TFSAAKFASFITQYSTGQVSVEIEWELQKNSKRWNPVEIQYTSNYKSNVVDFTVDTNVG 719

QY 721 YTEPRPIGTRYLTRPL 736
Db 720 YSEPRPIGTRYLTRNL 735

RESULT 8
AAU98974
ID AAU98974 standard; Protein; 735 AA.
XX

AC AAU98974;

XX 05-NOV-2002 (first entry)

XX Adeno-associated virus 2 (AAV2) vector, VP1 capsid protein.

XX Adeno-associated virus 2 vector; AAV2; cancer; VP1 capsid;

KW


```
CC oligopeptides may further be used to bind to host antibody to provide a
CC transient tolerant or non-responsive state.
XX
SQ Sequence 734 AA;

Query Match      84.6%; Score 3376; DB 22; Length 734;
Best Local Similarity 83.0%; Pred. No. 2e-262;
Matches 611; Conservative 51; Mismatches 72; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLEDNISEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLD 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEVPNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQ 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KGEVPNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ 120

QY 121 AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTDSE 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTDAD 180

QY 181 SVPDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGD RVI 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 SVPDPQPLGQPPAAPSGLGTTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVI 240

QY 241 TTSTRTWALPTYNNHLYKOISSASTGASNDNHYFGYSTPWGYDFDNRFCHFSPRDWQRL 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TTSTRTWALPTYNNHLYKOISSQS-GASNDNHYFGYSTPWGYDFDNRFCHFSPRDWQRL 299

QY 301 INNNWGFRPKRLNFKLFNIQKVEVTTNDGVTTIANNLTSTVQVFSDESQYLPVYLSAHQ 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 INNNWGFRPKRLNFKLFNIQKVEVTQNDGTTTIANNLTSTVQVETDSEYQLPYVLSAHQ 359

QY 361 GCLPPFPADVFMIPOYGYLTNNGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVP 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 GCLPPFPADVFVMPQYGYLTNNGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEDVP 419

QY 421 FHSSYAHQSQSLDRLMNP LIDQYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKNWLP 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 FHSSYAHQSQSLDRLMNP LIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLP 479

QY 481 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGV 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GPCYRQQRVSKTSADNNNSSEYTSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGV 539

QY 541 MIFGESAGASNTALDNVMTTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
   :|||: : : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 LIFGQSGSEKTNVDIEKVMITDEEIRTTN-VATEQYGSVSTNLQGRNQAAATADVNTQG 598

QY 601 ALPGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPVPANPPA 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 VLPGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPST 658

QY 661 EFSATKFAFITQYSTGQSVSEIEWELQENSKRWNPVQYTSNYAKSANVDFTVDNNGL 720
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 659 TFSAAKFAFITQYSTGQSVSEIEWELQENSKRWNPVQYTSNYKSNVNVDFTVDTNGV 718

QY 721 YTEPRPIGTRYLTRPL 736
   |:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 719 YSEPRPIGTRYLTRNL 734

RESULT 10
ID AAY71168
XX
AC AAY71168;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP2.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
```

```
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP2.
XX
OS Adeno associated virus serotype 1.
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00778.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host
XX
PS Claim 7; Page 93-95; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 cap protein VP2 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 599 AA;

Query Match      81.5%; Score 3251; DB 21; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.7e-252;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQPAKKRLNFGQTDSESVDPQPPLGEPPATPA 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQPAKKRLNFGQTDSESVDPQPPLGEPPATPA 60

QY 198 VGPTTMASGGAPMADNNEGADGVGNASGNWHCDSTWLGD RVIITSTRTWALPTYNNHLY 257
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 VGPTTMASGGAPMADNNEGADGVGNASGNWHCDSTWLGD RVIITSTRTWALPTYNNHLY 120

QY 258 KOISSASTGASNDNHYFGYSTPWGYDFDNRFCHFSPRDWQRLNNWGFPRKRLNFKLP 317
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 KOISSASTGASNDNHYFGYSTPWGYDFDNRFCHFSPRDWQRLNNWGFPRKRLNFKLP 180

QY 318 NIOQKEVTTNDGVTTIANNLSTVQVFSDESQYLPVYLSAHQGLPPFPADVFMIPOYG 377
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 NIOQKEVTTNDGVTTIANNLSTVQVFSDESQYLPVYLSAHQGLPPFPADVFMIPOYG 240

QY 378 YLTNNGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPFHSSYAHQSQSLDRLMNP 437
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 YLTNNGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPFHSSYAHQSQSLDRLMNP 300

QY 438 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 497
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360

QY 498 NSNFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGVMI FGKESAGASNTALDN 557
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 NSNFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGVMI FGKESAGASNTALDN 420
```


PT chimeric capsid having one non-native amino acid sequence and a desired
PT transgene -
XX
PS Disclosure; Page 51; 53pp; English.
XX
CC The invention provides a recombinant viral vector (RVV) comprising a
CC chimeric capsid (I) having at least one non-native amino acid sequence,
CC derived from a capsid protein domain of parvovirus (II), a virus (III),
CC or their combination, and a transgene flanked 5' and 3' by inverted
CC terminal repeat (ITR) sequences, derived from (II), (III), or their
CC combination. The RVV is useful for improving gene therapy in a subject
CC with a disorder, and for increasing the efficiency of entry into a cell,
CC which involves producing (I) encapsulating a viral vector, and contacting
CC a cell with RVV having (I) such that (I) binds to an attachment site on
CC the cell surface and permits the vector to enter the cell efficiently.
CC A pharmaceutical composition comprising RVV with (I) containing a
CC transgene sequence associated with a disease or a disorder such that
CC expression of the transgene would result in amelioration of the disease
CC or disorder such as inherited neurological and metabolic diseases e.g.
CC lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polyneuropathy,
CC Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood
CC e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic
CC fibrosis, diabetes, diseases associated with hormone deficiencies,
CC retinoblastoma and various types of neoplastic cells which include
CC tumours especially central nervous system tumours, neoplasms, carcinomas,
CC sarcomas, leukemias and lymphoma. The present sequence represents the
CC adeno-associated virus 2 (AAV-2) major coat protein VP2. AAV-2 sequences
CC are used in the construction of a chimeric vector.

XX Sequence 598 AA;

Query Match 69.2%; Score 2759.5; DB 22; Length 598;
Best Local Similarity 82.8%; Pred. No. 5.8e-213;
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy 139 APGKKRPVEQSPQEPDSSGIGKTGQQPAKKRLNFGQTGDSVPDPQPLGEPPTAAV 198
Db 2 APGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQLGQPPAASGL 61
Qy 199 GPTTMASGGGAPMADNNEGADGVGNASGNHWCDSWLGDRTTSTRTWALPTVNNHLYK 258
Db 62 GTNTMATGSGAPMADNNEGADGVGNSSGNHWCDSWTMGDRVITSTRTWALPTVNNHLYK 121
Qy 259 QISSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLNINNWGFRPKRLNFKLFN 318
Db 122 QISSQS-GASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLNINNWGFRPKRLNFKLFN 180
Qy 319 IQVKEVTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 378
Db 181 IQVKEVTQNDGTTTIANNLTSTVQVTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 240
Qy 379 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEDVPFHSSYAHQSOLDRLMNPL 438
Db 241 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEDVPFHSSYAHQSOLDRLMNPL 300
Qy 439 IDQYLYLNRNTQNSGSAQNKKLLFSGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNN 498
Db 301 IDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNN 360
Qy 499 SNFTWTGASKYNLNGRESIIINPGTAMASHKDEDEKFFPMSGVMIFFGESAGASNTALDNNV 558
Db 361 SEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEDEKFFPQSGVLIFGKQSEKTNVDIEKV 420
Qy 559 MITDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDHAMGALPGMWQDRDVLQGP 618
Db 421 MITDEEIRITNPVATEQYGSVSTNLQRNQAATADVNTQGLVPGMWQDRDVLQGP 480
Qy 619 WAKIPHTDGHFSPMLMGFGFLKXNPPQILIKNTVPANPPAEFSATKFAFITQYSTGQ 678
Db 481 WAKIPHTDGHFSPMLMGFGFLKXNPPQILIKNTVPANPSTTFAAKFAFITQYSTGQ 540
Qy 679 VSVEIEWELQENSKRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736

Db 541 VSVEIEWELQENSKRWNPVEIQYTSNYKSNVNVDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 13

AAM51509

ID AAM51509 standard; protein; 598 AA.

XX AAM51509;

DT 02-JAN-2002 (first entry)

XX Adeno-associated virus VP2 capsid protein.

XX Adeno-associated virus; AAV; VP2; capsid; virus-like particle;

KW nuclear localisation signal; VP3.

XX Adeno associated virus.

XX JP2001169777-A.

XX 26-JUN-2001.

XX 30-JUL-1999; 99JP-0249140.

XX 30-JUL-1999; 99JP-0249140.

XX (HAND/) HANDA H.

XX WPI; 2001-599854/68.

XX New virus-like particles from VP3 capsid protein of adeno-associated
PT virus, comprise a peptide containing a nucleus-shifting signal
PT connected to its N-terminal -

PS Claim 1; Page 14-16; 33pp; Japanese.

XX The present sequence is provided in a specification relating to a
CC virus-like particle-forming protein, and to a peptide containing a
CC nuclear-shifting signal at its N-terminus that can form virus-like
CC particles by shifting to the nucleus of the animal cell in which it
CC is expressed. The method is used for forming virus-like particles
CC from the VP3 capsid protein of adeno-associated virus (AAV).

XX Sequence 598 AA;

Query Match 69.2%; Score 2759.5; DB 22; Length 598;

Best Local Similarity 82.8%; Pred. No. 5.8e-213;

Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy 139 APGKKRPVEQSPQEPDSSGIGKTGQQPAKKRLNFGQTGDSVPDPQPLGEPPTAAV 198

Db 2 APGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQLGQPPAASGL 61

Qy 199 GPTTMASGGGAPMADNNEGADGVGNASGNHWCDSWLGDRTTSTRTWALPTVNNHLYK 258

Db 62 GTNTMATGSGAPMADNNEGADGVGNSSGNHWCDSWTMGDRVITSTRTWALPTVNNHLYK 121

Qy 259 QISSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLNINNWGFRPKRLNFKLFN 318

Db 122 QISSQS-GASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLNINNWGFRPKRLNFKLFN 180

Qy 319 IQVKEVTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 378

Db 181 IQVKEVTQNDGTTTIANNLTSTVQVTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 240

Qy 379 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEDVPFHSSYAHQSOLDRLMNPL 438

Db 241 LTLNNGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEDVPFHSSYAHQSOLDRLMNPL 300

Qy 439 IDQYLYLNRNTQNSGSAQNKKLLFSGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNN 498

Db 301 IDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNN 360

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:00:42 ; Search time 17.3269 Seconds
(without alignments)
4084.983 Million cell updates/sec

Title: US-09-807-802A-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177.5	54.6	504	1 VCPV3A	coat protein - ade
2	2168.5	54.4	732	2 S52210	coat protein VP1 -
3	798.5	20.0	673	1 VCPVB5	coat protein VP1 -
4	600.5	15.1	781	1 VCPV19	coat protein VP1 -
5	482	12.1	729	1 VCPVNA	coat protein VP1 -
6	482	12.1	729	1 A60006	coat protein VP1 -
7	444.5	11.1	727	1 VCPV1F	coat protein VP1 -
8	439.5	11.0	722	1 VCPVME	coat protein VP1 -
9	431.5	10.8	727	1 VCPVFP	coat protein VP1 -
10	426	10.7	723	1 VCPVPP	coat protein VP1 -
11	424.5	10.6	718	1 VCPVIM	coat protein VP1 -
12	421.5	10.6	748	1 VCPVCP	coat protein VP1 -
13	407.5	10.2	737	1 VCPVCD	coat protein VP1 -
14	404	10.1	722	1 VCPVV2	coat protein VP1 -
15	385.5	9.7	722	1 VCPVCN	coat protein VP1 -
16	377	9.5	716	1 VCPV2M	coat protein VP1 -
17	234	5.9	587	1 B44276	coat protein VP1 -
18	199.5	5.0	810	2 A44054	orf1 protein - Jun
19	198	5.0	584	2 S49594	capsid protein VP2
20	191	4.8	702	1 VCPVAP	coat protein VP1 -
21	143.5	3.6	648	2 S50856	whn protein - rat
22	139	3.5	931	2 T49710	related to glucan
23	131.5	3.3	1446	1 A45344	immediate-early pr
24	130.5	3.3	1145	2 T18235	transcription acti
25	130	3.3	1338	2 T30565	MAP kinase kinase
26	129	3.2	1072	2 A86827	hypothetical prote
27	124.5	3.1	2493	2 A55481	adenylate cyclase
28	124	3.1	1296	2 C82521	hemolysin-type cal
29	124	3.1	3078	2 T28432	variant-specific s

30 123.5 3.1 690 2 AB0124 probable TonB-depe
31 123.5 3.1 2529 2 B64635 toxin-like outer m
32 123 3.1 1142 2 T37455 enamelin precursor
33 122.5 3.1 880 1 SYBSVS valine-tRNA ligase
34 121 3.0 635 2 F96660 protein F2K11.10 [
35 121 3.0 1113 2 S28925 nuclear pore compl
36 121 3.0 1549 2 I48103 type VII collagen
37 120 3.0 1379 2 T37752 hypothetical serin
38 119.5 3.0 1180 2 B86719 hypothetical prote
39 119.5 3.0 1394 2 S60762 IGA-specific serin
40 119.5 3.0 2004 2 AC0314 probable membrane
41 119 3.0 667 2 A41311 transcription fact
42 118.5 3.0 2399 2 H71879 toxin-like outer m
43 118 3.0 739 2 T52289 probable transketo
44 118 3.0 956 2 T08144 myrosinase-binding
45 117.5 2.9 1076 1 A35622 nuclear pore prote

coat protein - adeno-associated virus type 2
C;Species: adeno-associated virus type 2
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Jul-1999
C;Accession: A03698
R;Srivastava, A.; Lusby, E.W.; Berns, K.I.
J. Virol. 45, 555-564, 1983
A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.
A;Reference number: A03694; MUID:83164299; PMID:6300419
A;Accession: A03698
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-504 <SRI>
A;Cross-references: EMBL:J01901; NID:g209616; PIDN:AAA42376.1; PID:g209621; EMBL:M12405;
C;Superfamily: adeno-associated virus coat protein
C;Keywords: coat protein

RESULT 1

VCPV3A

Query Match	54.6%	Score 2177.5	DB 1	Length 504
Best Local Similarity	82.5%	Pred. No. 1.3e-134		
Matches 392	Conservative 33	Mismatches 47	Indels 3	Gaps 2
QY	203	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS	262	
Dbb	1	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLYKQISS	60	
QY	263	ASTGASNDNHFGYSTPWGYDFDNRFHCHFSRDRWRLINNNWGRPKRLNFKLFNIQVK	322	
Dbb	61	QS-GASNDNHFGYSTPWGYDFDNRFHCHFSRDRWRLINNNWGRPKRLNFKLFNIQVK	119	
QY	323	EVTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN	382	
Dbb	120	EVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN	179	
QY	383	NGSQAVGRSSFYCYLPPSQMLRTGNNFTFSYTFEEVPEFHSSYAHSQSLDRLMNPIDQY	442	
Dbb	180	NGSQAVGRSSFYCYLPPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY	239	
QY	443	LYYLNRTQNQSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT	502	
Dbb	240	LYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQORVSKTSADNNNSEYS	299	
QY	503	WTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGVMIFGKESAGASNTALDNVMTD	562	
Dbb	300	WTGATKYHLNGRDSLVP--AMASHKDDDEKFFPQSGVLI FGKQSEKTNVIEKVMITD	357	
QY	563	EEEEKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQDRDVLQGPWAKI	622	
Dbb	358	EEIEGTTNPVATEQYGSVSTNLQRGNRQAATADVTQGVLPGMVWQDRDVLQGPWAKI	417	
QY	623	PHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKPFASFITQYSTG	677	

Db 421 TKPL--KLTHSWQTNRSGLPPLKLTTEPTTEGDQHPGTLPAANTRKGYHQITINNSYTEAT 478

Qy 468 PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMA 525

Db 479 ----AIRP-----AQVGYNTPYMNFYSNGGPF-----LTPIVPTADT 512

Qy 526 SHKODEDKFFPMGSMVIFGKESAGASNTALD---NVMITDEEEK--ATNP-----VAT 574

Db 513 QYNDDEPN-----GAIRFTMDYQHGHLTSSQELERYTFNPQSKCGRAPK 557

Qy 575 ERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVLQGPPIWA 620

Db 558 QQFNQQAPLNLENTNNGTLLPSDPIGGKSNMHFMNTLNTYGLPLTALNNTAPVFPNGQIWD 617

Qy 621 KIPHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPPAEFSA-TKFAFITQYST 676

Db 618 KELDTDLKPLRH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQQPRIITYSN 671

Qy 677 QOVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728

Db 672 FWKGTLTFTAKMRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712

RESULT 6

A60006

coat protein VP1 - porcine parvovirus (strain 90HS)

N;Contains: coat protein VP2

C;Species: porcine parvovirus

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993

C;Accession: A60006

R;Sakurai, M.; Nishimori, T.; Ushimi, C.; Nakajima, H.

Virus Res. 13, 79-86, 1989

A;Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.

A;Reference number: A60006; MUID:89319168; PMID:2750278

A;Accession: A60006

A;Molecule type: DNA

A;Residues: 1-729 <SAK>

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein; glycoprotein

F;151-729/Product: coat protein VP2 #status predicted <VP2>

F;172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 12.1%; Score 482; DB 1; Length 729;

Best Local Similarity 25.7%; Pred. No. 1.3e-23;

Matches 197; Conservative 112; Mismatches 314; Indels 144; Gaps 35;

Qy 43 RGLVLPGYKYLGPFGNLDKGEVPNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOF 102

Db 9 RGLTLPGYKYLPGNSLDQGEPTNPSDAAAKEHDEAYDKYIKSGKNPYFYFSAADEKFIK 68

Qy 103 RLQEDTSFGNLRGRAVFOAKKRVLLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGK- 161

Db 69 ETEHAKDYGGKIGHYFFRA-KRAFRP--KLSETDSPITSQPEVRRSPRKHPSKPPGKR 125

Qy 162 -----TGQOPAKKRLNFGQTGDSESVDPQPLGEPPATPAAVGPPTTMASGGGAPMA 212

Db 126 PAPRHIFINLAKKAKGTSTNSNSMSSENVQHNPNINAGTEL-SATGNESGGGGGGG--- 181

Qy 213 DNNEGADGVGNASGNWH--CDSTWLGD---RVITSTRTWALPTYNNHLYKQIS--SAST 265

Db 182 -GGRGAGGVGVSTGSFNNQTEFQYLGEGLVIRITAHASRLIHLNMPHEHETYKRIHVLNSES 240

Qy 266 GAS----NDNHFGYSTPWGYDFDNRFCHFSPRDQWRLLNNWGFRRPKRLNFKLFNIQV 321

Db 241 GVAGQMVQDDAHTQMTVPNSLIDANAWGVFNFPADWQLISNNMTNINLVSEFEQEIFNVVL 300

Qy 322 KEVT---TNDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVPMIPQYGY 378

Db 301 KTITESATSPPTKIYNNDLTASIMVALDNTNLTPTTPAAPRSETLGFYPWLPTKPTQYRY 360

Qy 379 L-----TLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNNF-TFYSYTFE 417

Db 361 YLSCTRNLPPTYTQSQQITDSIQTGLHSDIMFYTTIENAVPIHLLRTGDEFSTGIYHFD 420

Qy 418 EVPFHSSYAHSQSLDR-----LMNPLI--DOVLYYLNRQTQNSGSAQNKDLLPSRGS 467

Db 421 TKPL--KLTHSWQTNRSGLPPLKLTTEPTTEGDQHPGTLPAANTRKGYHQITMNNSYTEAT 478

Qy 468 PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMA 525

Db 479 ----AIRP-----AQVGYNTPYMNFYSNGGPF-----LTPIVPTADT 512

Qy 526 SHKODEDKFFPMGSMVIFGKESAGASNTALDNVMTDEEEKATNP-----VATERFGT 579

Db 513 QYNDDE---PNGAIRFTMGYQHGLTTS-----SQELERYTFNPQSKCGRAPKQGFNQ 562

Qy 580 VA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVLQGPPIWAKIPHT 625

Db 563 QAPLNLENTNNGTLLPSDPIGGKPNMHFMNTLNTYGLPLTALNNTAPVFPNGQIWDKELDT 622

Qy 626 D--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPPAEFSA-TKFAFITQYSTGQVS 681

Db 623 DLKPRLLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQQPRIITYSNFWWKG 676

Qy 682 EIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728

Db 677 TLFTAKMRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712

RESULT 7

VCPVIF

coat protein VP1 - feline panleukopenia virus

N;Contains: coat protein VP2

C;Species: feline panleukopenia virus, FPLV

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C;Accession: A03701

R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.

J. Virol. 55, 574-587, 1985

A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus.

A;Reference number: A03697; MUID:85265017; PMID:2991581

A;Accession: A03701

A;Molecule type: DNA

A;Residues: 1-727 <CAR>

A;Cross-references: EMBL:M10824; NID:G333474; PIDN:AAA47161.1; PID:G333476

C;Genetics:

A;Introns: 11/1

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.1%; Score 444.5; DB 1; Length 727;

Best Local Similarity 23.7%; Pred. No. 3.7e-21;

Matches 187; Conservative 117; Mismatches 295; Indels 189; Gaps 35;

Qy 43 RGLVLPGYKYLGPFGNLDKGEVPNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOF 102

Db 10 RGLVPPPGYKYLGPNGNSLDQGEPTNPSDAAAKEHDEAYAAAYLRSKKNPYLYFSPADQRFID 69

Qy 103 RLQEDTSFGNLRGRAVFOAKKRVLLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKT 162

Db 70 QTKDATDWGGKIGHYFFRAKKAIAPVLTDTDPHPSTSRPTK-PTKRSKPPPHIFINLAK- 127

Qy 163 GQOPAKKRLNFGQTGDSESV-----DPQPLGEPPA--TPAAVGPPTTMASGGGAPMADNNE 216

Db 128 -----KKKAGAGQVKRDNQAPMSDGAQVQPDGGQPAVRNERATSGNGSGGGG-----GG 176

Qy 217 GADGVGNASGNWH-----CDSTWLGDVRITSTRTWALPTYNNHLYKQI-----SS 262

Db 177 GSGGVGISTGTFFNQTEFKLENGW--EITANSSRLVHLNMPESSENYKRVVWVNNMDKTA 234

Qy 263 ASTGASNDNHFGYSTPWGYDFDNRFCHFSPRDQWRLLNNWGFRRPKRLNFKLFNIQVK 322

Db 235 VKGNMALDDTHVQIVTPWSLVDANAWGVWFNPGDWQLIVNTMSSELHLVSFEQEIFNVVLK 294

Qy 323 EVT---TNDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGCL-----PPPPA----- 368

Db	295	TVSESATQPPTKVYNNDLTASLMVALDSNNTMPFTPAAMRSETLGFYFWKPTIPTPWRY	354
Qy	369	---DVFMIPQGYLTLNNGSQAAGRS-----SFYCLE-YFPSQMLRTGNF-TFS	413
Db	355	FQWDRTLIPSH-----TGTSGTPTNIYHGTDPDVQFYTIENSVPVHLLRTGDEFATGT	408
Qy	414	YTFEEVPFHSSYAHQSQSLDRLMN--PLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGM	471
Db	409	FPFDCKP--CRLTHWTQTNRALGLPP-----FLNSLPQSEGATNFGDIGVQQDKRRGV	459
Qy	472	S-----VQPKN-WLPGPCYRQORVSK-----TKTDNNSNFTWT	504
Db	460	TQMGNTDYITEATIMRPAEVGYSAPYYSFEASTQGPFKIPIAAGRGGAGTQDENQAA--D	516
Qy	505	GASKY---NLNGRESIINPGT-----AMASHKDDDEDKFFPMSGVMIFGKESAGASNTALD-	556
Db	517	GDPRYAFGRHQHGQKTTTGETPERFTYIAHQDT-----GRYPAGDWIQNINF	563
Qy	557	NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQ	616
Db	564	NLPVTNDNVLLPTDPIG---GKTGINY--TNIFNTYGPLTALNNVP-----PVYPNG	610
Qy	617	PIWAKIPHTDGHFHPSPMLMGFGFLK-----NPPPQILIKNTPVPAN---PPA	660
Db	611	QIWDKEFDTD-----LKPRLHVNAFPVCQNNCPGQLFVKVAPNLITNEYDPDA	657
Qy	661	EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNGL	720
Db	658	SANMSR---IVTYSDFWVKGLVFKAKLRASHWTWNPIQOMSIN-----VDNQFN	703
Qy	721	YTEPRPIG 728	
Db	704	YL-PNNIG 710	
RESULT 8			
VCPVME			
coat protein VP1 - mink enteritis virus (strain Abashiri)			
N:Contains: coat protein VP2			
C:Species: mink enteritis virus, MEV			
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Apr-1996			
C:Accession: B38350			
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishigurio, N.; Goto, H.; Shinag			
J. Gen. Virol. 72, 867-875, 1991			
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the			
A:Reference number: A38350; MUID:91202123; PMID:2016597			
A:Accession: B38350			
A:Molecule type: DNA			
A:Residues: 1-722 <KAR>			
A:Cross-references: GB:D00765			
C:Superfamily: parvovirus coat protein			
C:Keywords: coat protein			
P:139-722/Product: coat protein VP2 #status predicted <VP2>			
Query Match 11.0%; Score 439.5; DB 1; Length 722;			
Best Local Similarity 23.6%; Pred. No. 7.8e-21;			
Matches 186; Conservative 117; Mismatches 295; Indels 189; Gaps 35;			
Qy	44	GLVLPGYKYLGPFGNLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQER	103
Db	6	GLVPPGYKYLPGNSLDQGEPTNPSDAAAKEHDEAYAYALRSGKNPYLYFSPADQRFIDQ	65
Qy	104	LQEDTSFGGNLGRAVFAQKRVLEPLGLVEEGAKTAPGKKRPVEQSPQSPDSSSGIGKTG	163
Db	66	TKDATDWGGKIGHYFFRAKKAIAPVLTDTDPNPSTSRPTK-PTKRSKPPPHIFINLAK--	122
Qy	164	QOPAKKRLNFGQTDSESVP---DPQPLGEPPA--TPAAVGPTTMASSGGGAPMADNNEG	217
Db	123	---KKKAGAGQVKCRDNLAPMSDGAQVDPGQPAVRNERATGSGNGSGGGG-----GGG	172
Qy	218	ADGVGNASGNWH-----CDSTWLGDVITTTSTRTWALFTYNNHLYKQI-----SSA	263
Db	173	SGGVGISTGTFFNNQTEFKFLENGWV--EITANSSRLVHLNMPESENYKRVVNVNMDKTAV	230

Qy	264	STGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVKE	323
Db	231	KGNMALDDTHVQIVTPWSLVDANAWGVWFPNGDWQLIVNTMSELHVSFEQEIFNVVLKT	290
Qy	324	VT---TNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGCL-----PPPPA-----	368
Db	291	VSESATQPPTKVYNNDLTASLMVALDSNNTMPFTPAAMRSETLGFYPWKPTIPTPWRYYP	350
Qy	369	--DVFMIPQGYLTLNNGSQAAGRS-----SFYCLE-YFPSQMLRTGNF-TFSY	414
Db	351	QWDRTLIPSH-----TGTSGTPTNIYHGTDPDVQFYTIENSVPVHLLRTGDEFATGTF	404
Qy	415	TFEVFPFHSSYAHQSQSLDRLMN--PLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMS	472
Db	405	FFDCKP--CRLTHWTQTNRALGLPP-----FLNSLPQSEGATNFGDIGVQQDKRRGVT	455
Qy	473	-----VQPKN-WLPGPCYRQORVSK-----TKTDNNSNFTWTG	505
Db	456	QMGNTDYITEATIMRPAEVGYSAPYYSFEASTQGPFKTPIAAGRGGAGTQDENQAA---DG	512
Qy	506	ASKY---NLNGRESIINPGT-----AMASHKDDDEDKFFPMSGVMIFGKESAGASNTALD-N	557
Db	513	DPRYAFGRHQHGQKTTTGETPERFTYIAHQDT-----GRYPAGDWIQNINFN	559
Qy	558	VMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGP	617
Db	560	LPVTNDNVLLPTDPIG---GKTGINY--TNIFNTYGPLTALNNVP-----PVYPNGQ	606
Qy	618	IWAKIPHTDGHFHPSPMLMGFGFLK-----NPPPQILIKNTPVPAN---PPAE	661
Db	607	IWDKEFDTD-----LKPRLHVNAFPVCQNNCPGQLFVKVAPNLITNEYDPDAS	653
Qy	662	FSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNGLY	721
Db	654	ANMSR---IVTYSDFWVKGLVFKAKLRASHWTWNPIQOMSIN-----VDNQFNY	699
Qy	722	TEPRPIG 728	
Db	700	L-PNNIG 705	
RESULT 9			
VCPVPP			
coat protein VP1 - feline panleukopenia virus (strain 193)			
N:Contains: coat protein VP2			
C:Species: feline panleukopenia virus, FPLV			
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 12-Apr-1996			
C:Accession: B36608			
R:Martyn, J.C.; Davidson, B.E.; Studdert, M.J.			
J. Gen. Virol. 71, 2747-2753, 1990			
A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvo			
A:Reference number: A36608; MUID:91073139; PMID:2174965			
A:Accession: B36608			
A:Molecule type: DNA			
A:Residues: 1-727 <MAR>			
A:Cross-references: GB:X55115			
C:Superfamily: parvovirus coat protein			
C:Keywords: coat protein			
F:144-727/Product: coat protein VP2 #status predicted <VP2>			
Query Match 10.8%; Score 431.5; DB 1; Length 727;			
Best Local Similarity 23.5%; Pred. No. 2.6e-20;			
Matches 185; Conservative 117; Mismatches 297; Indels 189; Gaps 35;			
Qy	43	RLVLPGYKYLGPFGNLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOF	102
Db	10	RLVPPGYKYLPGNSLDQGEPTNPSDAAAKEHDEAYAAVYLRSGKNPYLYFSPADQRFID	69
Qy	103	RLQEDTSFGGNLGRAVFAQKRVLEPLGLVEEGAKTAPGKKRPVEQSPQSPDSSSGIGKT	162
Db	70	QTKDAKDWGGKIGHYFFRAKKAIAPVLTDTDPHPSTSRPTK-PTKRSKPPPHIFINLAK-	127

QY 163 GQPAKRLNFGQTGDSVSP---DQPLGEPPA--TPAAVGPTTMSGGGAPMADNNE 216
Db 128 -----KKKAGAGQVKRDNLAPMSDGAQPDGGQPAVRNERATGSGNGSGGG-----GG 176
QY 217 GADGVGNASGNWH-----CDSTWLGDVRITTTSTRTWALPTYNNHLYKQI-----SS 262
Db 177 GSGGVGISTGTENNQTEFKLENGWV--EITANSSRLVHLNMPESENYKRVVYNNMDKTA 234
QY 263 ASTGASNDNHYFGYSTPWGYPDFNRFCHFSPRDWQRLNINNWFPRKRLNFKLFNIQVK 322
Db 235 VKGNMALDDIHQIVTPWSLVDANAGVWFNPGDWQLIVNTMSELHLVSEFELEFNVVLK 294
QY 323 EVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGL-----PPPPA----- 368
Db 295 TVSESATQPPTKVYNNDLTASLMVALDSNNTMPFTPAAWRSETLGFYPWKPTIPTPWRY 354
QY 369 ---DVFMIPQGYLTLNNGSOAVGRS-----SFYCLE-YFPQSOMLRTGNMF-TFS 413
Db 355 FQWDRTLIPSH-----TGTSGTPTNVYHGTDPDDVQFYTIENSVPVHLLRTGDEFATGT 408
QY 414 YTFEEVPFHSSYAHQSOLDRLMN--PLIDQLYLYLNRTQNGSGSAQNKOLLFSRGS 471
Db 409 FPFDCKP--CRLTHTWTQNRALGLPP-----FLNSLPQSEGATNFGDIGVQDCKRRGV 459
QY 472 S-----VQPKN-WLPGPCYRQQRVSK-----TKTDNNSNFTWT 504
Db 460 TQMGNTDYITEATIMRPAEVGYSAPYYSFEASTQGPFKTPIAAGRGAQTDENOAA---D 516
QY 505 GASKY---NLNGRESIINPGT---AMASHKDDDEKFFPMSGVMIFGKESAGASNTALD- 556
Db 517 GDPRYAFGRHQGKTTTGTETPERFTYIAHQDT-----GRYPEGDWIQNINF 563
QY 557 NVMITDEEEIKATNPVATEREFTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVYLO 616
Db 564 NLPVTNDNVLLPTDPIG-----GKTGINY--TNIFNTYGLTALNNVP-----PVYPNG 610
QY 617 PIWAKIPHTDGHFHPSPMLMGFGGLK-----NPPQILIKNTVPVPAN---PPA 660
Db 611 QIWDKEBFTD-----LKPRLHVNAFVVCQNNCPGQLFVKVAPNLTNEDPD 657
QY 661 EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDN 720
Db 658 SANMSR-----IVTYSDFWFKGLVEKAKLRASHTWNPIQQMSIN-----VDNQFN 703
QY 721 YTEPRPIG 728
Db 704 YV-PNNIG 710

RESULT 10
VCPVPP
coat protein VP1 - porcine parvovirus
N;Contains: coat protein VP2
C;Species: porcine parvovirus
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 12-Apr-1996
C;Accession: B33302
R;Ranz, A.I.; Mancus, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A;Title: Porcine parvovirus: DNA sequence and genome organization.
A;Reference number: A33302; MUID:90010964; PMID:2794971
A;Accession: B33302
A;Molecule type: DNA
A;Residues: 1-723 <RAN>
A;Cross-references: EMBL:D00623
C;Genetics:
A;Introns: 10/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.7%; Score 426; DB 1; Length 723;
Best Local Similarity 24.9%; Pred. No. 6e-20;
Matches 190; Conservative 109; Mismatches 308; Indels 156; Gaps 36;

QY 53 LGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGG 112
Db 13 LGPGNSLDQGEPTNPSDAAAKEHDEAYDKYIKSGKNPTFFYPSAADEKFIKETEHAKDYGG 72
QY 113 NLGRAVFOAKKRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGK-----T 162
Db 73 KIGHYFFRAKRAFAPKL---SETDSPTSQQPEVRRSPRKHPGSKPPKRPAPRHIFINL 129
QY 163 GQPAKRLNFGQTGDSVSPDPQPLGEPPATP-AAVGPTTMSGGGAPMADNNEGADGV 221
Db 130 AKKAKGTSNTNSMSSENVEQHNPIN--AATELSATGNESGGGGGG-----GGRGAGGV 183
QY 222 GNASGNWH--CDSTWLGD---RVITTTSTRTWALPTYNNHLYKQI-----SSASTGAS-ND 270
Db 184 GVSTGSFNNQTEFYQLGEGLVRIATAHASRLIHLNMPHEHYKRIHVLNSESAGOMVQD 243
QY 271 NHYFGYSTPWGYPDFNRFCHFSPRDWQRLNINNWFPRKRLNFKLFNIQVKEVT---TN 327
Db 244 DAHTQMVTPWSLIDANAGVWFNPADWQLISNNMTEINLVSFEQAIFNVVLKTITESATS 303
QY 328 DGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIQYGYL----- 379
Db 304 PPTKIYNNDLTASLMVALDTNNTLPTTPAAPRSETLGFYPWLPTKPTQYRYLSCIRNLN 363
QY 380 --TLNNGSOAVGRSS-----FYCLE-YFPQSOMLRTGNMF-TFSYTFEEVPFHSSYA 426
Db 364 PPTYTGQSQPNRLNTRLHSDIMFYTIENAVPIHLRLTGDEFSTGIYHFDTKPL--KLT 421
QY 427 HSQSLDR-----LMNPLI--DQLYLYLNRTQNGSGSAQNKOLLFSRGS 476
Db 422 HSWQTNRSGLPPLKLLTEPTTEGQHPGTLPAANTRKGYHOTINNSYTEAT---AIRP- 476
QY 477 NWLPGPCYRQQRVSKTKTDNNSNFTWTGASKYNLNGRESIINP--GTAMASHKDDDEK 534
Db 477 -----AQVGYNTPYMNFEYSNGGPF-----LTPIVPTADTQYNDDEPN- 514
QY 535 FPMMSGVMIFGKESAGASNTALD---NVMITDEEEIK--ATNP-----VATERFGTVA-V 582
Db 515 -----GAIRFTMDYQHGHLTSSQELERYTFENPQSKCGRAPKQFNNQQA 560
QY 583 NFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVYLOQPIWAKIPHTD--G 627
Db 561 NLENTNNGTLLPSDPIGGKSNMHFMNTLNTYGPLTALNNTAPVFPNGQIWDKELDTDLKP 620
QY 628 HFHPSPLMGFGGLK-NPPPQILIKNTVPVPANPPAPFSA-TKFASFITQYSTGQVSVEIEW 685
Db 621 RLH---VTAPFVCKNPNPPGQLFVKIAP---NLTDDEFNADSPQOPRIITYSNFWWKGLTLTF 674
QY 686 ELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728
Db 675 TAKMRSSNMWNPIQOHTT-----TAENIGKYI-PTNIG 706

RESULT 11
VCPVIM

coat protein VP1 - minute virus of mice (strain MVMi)
C;Species: minute virus of mice, murine parvovirus
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 28-Jul-2000
C;Accession: B23008; B29510
R;Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A;Title: DNA sequence comparison between two tissue-specific variants of the autonomous I
A;Reference number: A23008; MUID:85242059; PMID:3855242
A;Accession: B23008
A;Molecule type: DNA
A;Residues: 1-718 <SAH>
A;Cross-references: EMBL:X02481; NID:G60918; PIDN:CAB46507.1; PID:G5419928
R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A;Title: DNA sequence of the lymphotropic variant of minute virus of mice, MVM(1), and C
A;Reference number: A29510; MUID:86115415; PMID:3502703
A;Accession: B29510

VCPVCD
coat protein VP1 - canine parvovirus (strain CPV-d)
N;Contains: coat protein VP2
C;Species: canine parvovirus, CPV
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A31163
R;Parrish, C.R.; Aquadro, C.F.; Carmichael, L.E.
Virology 166, 293-307, 1988
A;Title: Canine host range and a specific epitope map along with variant sequences in the
A;Reference number: A31163; MUID:89020796; PMID:3176341
A;Accession: A31163
A;Molecule type: DNA
A;Residues: 1-737 <PAR>
A;Cross-references: EMBL:M23255; NID:g3333467; PIDN:AAA47158.1; PID:g3333468
C;Genetics:
A;Introns: 26/3
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.2%; Score 407.5; DB 1; Length 737;
Best Local Similarity 22.8%; Pred. No. 9.9e-19;
Matches 181; Conservative 112; Mismatches 301; Indels 201; Gaps 32;

QY 42 GRGLVLPGYKYLPGFNGLDKGEVPNAADAAAEHDKAYDQQLKAGDNPYLRYNHADAEEFQ 101
Db 19 GEGKDLITYKYLPGNSLDQGEPTNPSDAAAKEHDEAYAAALRSCKNPYLIFSPADQRFI 78
QY 102 ERLQEDTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGK 161
Db 79 DQTKDAKDWGGKIGHYFFRAKKAIAPLVLTDPDHPSTSRPTK-PTKRSKPPPHIPINLAK 137
QY 162 TGOQPAKRLNFGQTGDSVP----DPQLGEPPA--TPAAVGPTTMASSGGGAPMADNN 215
Db 138 -----KKKAGAGQVKRDLNAPMSDGAVPDGGQPAVRNENATSGNGSGGGG-----G 185
QY 216 EGADGVGNASGNWH-----CDSTWLGDRVITTTSTRTWALPTYNNHLYKQI-----S 261
Db 186 GSGGVGISTGTFFNQTEFKFLBNGWV--EITANSSRLVHLNMPESENYRRVVVNNMDKT 243
QY 262 SASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNNGWFRPKRLNFKLFIQV 321
Db 244 AVNGNMALDDIHAQIVTPWSLVDANAGWVFNPGDWQLIVNTMSLHVSFEQEIFNVVL 303
QY 322 KEVT---TNDGVTTIANNLSTVQVSDSEYQLPYVLGSAHQCL-----PPFPA---- 368
Db 304 KTVSESATQPTKVVNNDLTASLMVALDSNNTMPFTPAARSETLGFYPWKPTIPTPWRY 363
QY 369 ----DVFMPQYGYLTLLNGSQAVERS-----SFYCLE-YFPSQMLRTGNFF-TF 412
Db 364 YFQWDRLLPSH-----TGTSGETPTNIYHGTDPDDVQFYTIENSVPVHLLRTGDEFATG 417
QY 413 SYTFEEVFPFHSSYAHQSQSLDRLMN--PLIDQYLYLNRNTQNSGSAQNKDLFSRGSAG 470
Db 418 TFFFDCKP--CRLTHTWQTNRALGLPP-----FLNSLPQSEGATNFGDI-----G 460
QY 471 MSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF-----TWFGASKYNLNGRESIINP- 520
Db 461 V-----QQDKRRGVQTQMGNTNYITEATIMREAEGVGSAPYYSFEASTQGP 506
QY 521 ----GTAMASHKDDKEDKFFPMGVMIFGKESAGASNTALD----- 556
Db 507 KTIPIAAGRGGAQTDENQAADGNPRYAFGRHQGKTTTGTETPERFTYIAHQDTGRYPEGD 566
QY 557 -----NVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATDVHAMGALPGMVWD 609
Db 567 WIQINFNLPVTDNVLPTDPIG-----GKTGINY--TNIFNTYGPLTALNNVP----- 614
QY 610 RDVYLGPIWAKIPHTDGHFHPSPMLGCGFLK-----NPPPIILIKNTPVPA 656
Db 615 -PVYPNGQIWDKEFDTD-----LKPRLHVNAPFVQCQNCPCGQLFVKVAPNLT 660
QY 657 N---PPAEFSATKPASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDF 713

Db 661 NEYDPDASANMSR-----IVTYSDFWMKGLVFKAKLRASHTWNPIQQMSIN----- 707
QY 714 TVDNNGLYTEPRPIG 728
Db 708 -VDNQFNXYV-PSNIG 720

RESULT 14
VCPVV2
coat protein VP1 - parvovirus H1
C;Species: parvovirus H1
A;Note: host Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Apr-1994
C;Accession: A03699
R;Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid
A;Reference number: A03695; MUID:83112183; PMID:6823009
A;Accession: A03699
A;Molecule type: DNA
A;Residues: 1-722 <RHO>
A;Cross-references: EMBL:X01457; EMBL:J02198
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match 10.1%; Score 404; DB 1; Length 722;
Best Local Similarity 23.3%; Pred. No. 1.6e-18;
Matches 175; Conservative 108; Mismatches 290; Indels 178; Gaps 30;

QY 57 NGLDKGEPVNAADAAAEHDKAYDQQLKAGDNPYLRYNHADAEEFQERLQEDTSFGGNLGR 116
Db 12 NSLDQGEPTNPSDAAAKEHDEAYDYIKSGKNPYLYFSPADQRFIDQTKDAKDWGGKVG 71
QY 117 AVFOAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQ---PA----- 167
Db 72 YFFRTTKRAFAPKLTDS-----PG-----TSGVSRPGKRTKPPAHIFVNO 112
QY 168 ----KKRLNFGQTGDSVPDPQPLGEPPATPAAVGPTTMASSGGGAPMADNNEGADGVGN 223
Db 113 ARAKKRLASLAAQORTLTMSDGTETNPDTGIANARVERSADGGG-SGGGSGGGGIGV 171
QY 224 ASGNWHCDSTW--LGDRVITTTSTRTWAL-----PTYNHLYKQISSASTGA 267
Db 172 STGTYDNQTTYKFLGDGWVEITAHASRLHLGMPSPSENYCRVTVHNNQTTGHGTVKGNM 231
QY 268 SNDNHYFGYSTPWGYFDNRFHCHFSPRDWQRLINNNGWFRPKRLNFKLFIQVKEVTT- 326
Db 232 AYDTHQQIW-TPWSLVDANAGWVFPQSDWQFIQNSMESLNLDLSQELFNVVVKTVTTEQ 290
QY 327 ----NDGVTTIANNLSTVQVSDSEYQLPYVLGSAHQCL-----PPFPAD---VFMI 373
Db 291 QGAGQDAIKVYNNDLTACMVVALDSNNILPTPAAQTSSETLGFYPWKPTAPAPRYFFFM 350
QY 374 POYGYLTLLNGSQ-----AVG-----RSSFYCLE-YFPSQMLRTGNNTFF-SYTPBEVP 420
Db 351 PRQLSVTSSNSAEGTQITDTIGEPOALNSQFTIENTLPITLLRTGDEFTTGTIYFNTDP 410
QY 421 FHSSYA-----HSQSLDRLMN-PLIDQYLYLNRNTQNSGSAQNKOLLF----- 463
Db 411 LKLTHTWQTNRHLCQLQGITDLPTSDTATASLTANGDRFGSTQTNVNVYVTEALRTRPAQ 470
QY 464 -----SRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNL 511
Db 471 IGFMQPHDNFEANRGGFVKVPVP-----LDITAGEDHDAN----GAIRFNY 513
QY 512 NGR--ESIINPGTA-----MASHKDDKEDKFFPMGVMIFGKESAGASNTALDNVM 559
Db 514 GKQHGEDWAKQGAAPERYTWDAIDSAAGRDTARCFV-----QSAPISIPPNQNI 563
QY 560 ITDEEEIKA-TNPVATERFGTVAVNFQSSSTDPAATDVHAMGALPGMVWDQDRDVYLGQPI 618
Db 564 LOREDAIAGRTNMHYTNVFNYSYGLPSAFPHDP-----IYPNGQI 603

Db 240 LYYLSRTNTPSGTTQSRLLQFSQAGASDIRDOSRNWLPGPCCYRQORVSKTSADNNNSEYS 299

QY 503 WTGASKYNLNGRESIINPGTAMASHKDDKFFPMGVMIFGKESAGASNTALDNVMTD 562

Db 300 WTGATKYHLNGRDSLVP--AMASHKDDKFFPMGVMIFGKESAGASNTALDNVMTD 357

QY 563 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPWAKI 622

Db 358 EEEIGTTNPVATEQYGSVSTNLQRNGRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 417

QY 623 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTFQYSTG 677

Db 418 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASTFQYSTG 472

RESULT 2

COAT_PAVBO STANDARD; PRT; 673 AA.

AC P07297; Q84374;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Coat protein VP2 [Contains: Coat protein VP3].

OS Bovine parvovirus (BPV).

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OX NCBI_TaxID=10784;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87061184; PubMed=3783814;

RA Chen K.C., Shull B.C., Moses E.A., Lederman M., Stout E.R.,

RA Bates R.C.;

RT "Complete nucleotide sequence and genome organization of bovine

RT parvovirus.";

RL J. Virol. 60:1085-1097(1986).

CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; M14363; AAB59848.1; -.

DR EMBL; M14363; AAB59849.1; -.

DR PIR; A26104; VCPVB5.

DR HSSP; P30129; 4DPV.

DR InterPro; IPR001403; Parvo coat.

DR Pfam; PF00740; Parvo_coat; 1.

KW Coat protein; Glycoprotein.

FT CHAIN 1 673 COAT PROTEIN VP2.

FT CHAIN 138 673 COAT PROTEIN VP3.

FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DOMAIN 163 187 GLY-RICH.

SQ SEQUENCE 673 AA; 75103 MW; 5F244642B2214831 CRC64;

Query Match 20.0%; Score 798.5; DB 1; Length 673;

Best Local Similarity 28.6%; Pred. No. 9.1e-42;

Matches 214; Conservative 105; Mismatches 275; Indels 155; Gaps 21;

QY 30 PKPKANQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNP 89

Db 3 PTNKANSKK-----GLTLPGNYLGPFGNLSFAGAPVNAKADAAARKHDFGYSDLKEGKNP 57

QY 90 YLRYNHADAEEFOERLQEDTSFGNGLGRAVFOAKRVLPLEGLVEEGAKTAPGKRPEQS 149

Db 58 YLYFNTHDQNLIDELKDDTSFGGKLARGVFQIKKALAPAL-----PGTSKGGDRALKRK 111

QY 150 PQEPDSSSGIGKTGQQAQKRLNFGQTGDSSEVPDPQLGEP--PATPAAVGPTTMASSG 207

Db 112 LYFARSNKGAKKANREPAPSTSNQNNMEVSNNDIPNDEAGNQPIELATRSVSGSVGGG 171

QY 208 GAPMADNNEGADGVGNASGNWCHDSTWLGDRTTSTRTWALPTYNHLYKQISSASTGA 267

Db 172 -----RGGSGVGYSTGGTGGTIFSENIIVTKTRQICDIKNGHLYKS-EVLNTGD 222

QY 268 SNDNHYFGYSTPGYFDFNRHCHFSRDPQWRLNNNNGWFRPKRLNFKLFNIQKVEVTN 327

Db 223 TAHRQY-AITTPWSYFNQYSSHSFSPNDWQHLVNDYERFRPKAMIVRVNLIQIKQIMTD 281

QY 328 DGVTTIANN-LTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTL----- 381

Db 282 GAMGTVYNNDLTAGMHIFCDGHDHRYPYVQHPWDDQCMPELPSIWELPQYAYIPAPISVV 341

QY 382 -NNGSQAVGRS-----SFYCLEYPPSQMLRTGNFTFSYTFEEVPHSSYAHSQSLDRLM 435

Db 342 DNNTTNTVEHLLKGVPLYMLENSDHEVLRNG-----RIV 376

QY 436 NPLIDQYLYLNRTQN-----QSGSAQNKOLLFSRGS-----AGMSVQPKNWLP 480

Db 377 RIYIQLWRLMRDRKQHHIQHASDDVQSTGQKQKNLLIQRKQPNKQRFQNAALRTSNWMS 436

QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMG 540

Db 437 GP-----GIARGTHNATLQTSAGALVTMTNGAD-----VSGV 470

QY 541 M-----IFGKE-----SAGASNTALDNVMTDEEEIKATNPVATE 575

Db 471 RAVRVGYSTDPIYGGQQPESDLLRLRYASAAEGQONPILN-----AARH 516

QY 576 RFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKIPHTDGHFHPSPLM 635

Db 517 TFTREARTKLITGSGNGADGDYKEMWMLPNQMWDSAPISRYNPPIWVPRVNRKTLTDTQD 576

QY 636 GGFGLKNPPQILIK--NTPVPANPPAEFSATKFASTFQYSTGQVSVEIEWELQKENS 693

Db 577 GSIPMSHPPGTIFIKLARIPVPGND-----SFLNIYVTGQVSCVVEVEKRGTK 627

QY 694 RWNPEVQYTSNYAKSANVD-FTVDNNGLY 721

Db 628 NWRPEYMHs---ATNMSVDAYTINNAGVY 653

RESULT 3

COAT_PAVHB STANDARD; PRT; 781 AA.

AC P07299;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-APR-1990 (Rel. 14, Last annotation update)

DE Probable coat protein VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Isolate AU;

RX MEDLINE=86200451; PubMed=3701931;

RA Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;

RT "Nucleotide sequence and genome organization of human parvovirus B19

RT isolated from the serum of a child during aplastic crisis.";

RL J. Virol. 58:921-936(1986).

CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR	EMBL; M13178; AAA66867.1; -	
DR	PIR; A24299; VCPV19.	
DR	InterPro; IPR001403; Parvo_coat.	
DR	Pfam; PF00740; Parvo_coat; 1.	
KW	Coat protein; Glycoprotein.	
FT	CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE 781 AA; 86015 MW; 8C6254DBD0576B07 CRC64;	
Query Match 15.1%; Score 600.5; DB 1; Length 781;		
Best Local Similarity 26.6%; Pred. No. 1.6e-29;		
Matches 182; Conservative 95; Mismatches 300; Indels 107; Gaps 20;		
Qy	47 LPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQRLQE 106	
Db	125 LPGTNVYVGPNGNELQAGPPQSAVDASAARIHDFRYSQLAKLGINPYTHWTVADEELLKNIK 184	
Qy	107 DTSPGNGLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQQP 166	
Db	185 ETGFAQVVKDYF-----TLKGAAPVAHFQ-----GSLPEVP 217	
Qy	167 AKKRLNFGQTGDSSEVPDPQPLGEPPTAPAAVGPTTMASGSGAPMADNNEGADGVGNASG 226	
Db	218 AYNA-----SEKYPMTSVNSAEASTGA-----GGG-----GNSVKS 250	
Qy	227 NWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISSASTGASND-----NHYFGY 276	
Db	251 MWSEGATFSANSVCTFSRQLIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGY 310	
Qy	277 STPMGYDFNRFHCHFSPRDWRLNNWGRFPRKRLNFKLENIQVKEVT--TNDGVTTIA 334	
Db	311 STPWRYLDFNALNLFSPLEFOHLIENYGSIAPDALTVTISEIAVKDVTDKTGGV-QVT 369	
Qy	335 NNLTSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPOYGYLTNN-GSQAVG---- 389	
Db	370 DSTTGRCLMLVDHEYKYPYVLGGQDTLAPELPIWVYFPQYAYLTVGDVNTQGISGDSK 429	
Qy	390 -----RSSFYCLEYFSPQMLRTGNNTFFSYTFEEVPPFHSSVAHSQSLDRLMNLIDQYLY 444	
Db	430 KLASSEAFYVLEHSSFQLLTGCTASMSYKFPVPVPPENLEGCSQHFYEMYNPL---YGS 486	
Qy	445 YLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCY3QQRVSKTKTDNNNSNFTWT 504	
Db	487 RLGVDPDTLGGDPKFRSL-----THEDHAIQPNFMPGPLVNSVSTKEGDSSTGACKALT 541	
Qy	505 GASKYNLNGRESIINPG-TAMASHKDDDEDKFFPMMSGVMIFGKESAGASNTALDNV----- 558	
Db	542 GLSTGTSQNTRI5LRPGVPVQPYHHWDTDKYVTGAINAISHCQTTYG---NAEDKEYQQGV 598	
Qy	559 --MITDEEEIKATNPVATERFGTVAVNFQSSSTDPA2GDVHAMGALPGMVWQDRDYLQG 616	
Db	599 GRFPNEKEQLKOLQGLNMHTY-----FPNKG7QYTDQIE-RPLMVGSVMNRRALHYES 651	
Qy	617 PIWAKIPHTDGHFHS-PLMGFGFLKNPPPPQILIKNTPVPANPPAEFSATKFASFITQYS 675	
Db	652 QLWSKIPNLDDSFKTQFAALGGWGLHQPPPPQIFLK--ILPQSGPIGGIKSMGITTILVQYA 709	
Qy	676 TGQVSVEIEWEL-QKENSKRWNPE 698	
Db	710 VGIMVTWMTFKLGPVKATGRWNPQ 733	
RESULT 4		
COAL_PAVP9 STANDARD; PRT; 729 AA.		
AC	P33484;	
DT	01-FEB-1994 (Rel. 28, Created)	
DT	01-FEB-1994 (Rel. 28, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Coat protein Vp1 [Contains: Coat protein VP2].	
OS	Porcine parvovirus (strain 90HS) (PPV).	

OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	
OX	NCBI_TaxID=33725;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=89319168; PubMed=2750278;	
RA	Sakurai M., Nishimori T., Ushimi C., Nakajima H.;	
RT	"Nucleotide sequence of capsid protein gene of porcine parvovirus.";	
RL	Virus Res. 13:79-86(1989).	
CC	- - SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED	
CC	VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.	
CC	- - MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT	
CC	OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.	
CC	- - SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.	
DR	PIR; A60006; A60006.	
DR	HSSP; P30129; 4DPV.	
DR	InterPro; IPR001403; Parvo_coat.	
DR	Pfam; PF00740; Parvo_coat; 1.	
KW	Coat protein; Glycoprotein.	
FT	CHAIN 1 729 COAT PROTEIN VP1.	
FT	CHAIN 151 729 COAT PROTEIN VP2.	
FT	CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE 729 AA; 80938 MW; 92538BFF9A0C78E6 CRC64;	
Query Match 12.1%; Score 482; DB 1; Length 729;		
Best Local Similarity 25.7%; Pred. No. 2.8e-22;		
Matches 197; Conservative 112; Mismatches 314; Indels 144; Gaps 35;		
Qy	43 RGLVLPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQ 102	
Db	9 RGLTLPGYKYLGPNGSLDQGEPTNP5DAAAKEHDEAYDKYIKSGKNPYFYPSAADEKFIK 68	
Qy	103 RLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGK- 161	
Db	69 ETEHAKDYGGKIGHYFFRA-KRAFRP--KLSETDSPTTSQQPEVRRSPRKHPGSKPPGKR 125	
Qy	162 -----TGQOPAKKRLNFGQTGDSSEVPDPQPLGEPPTAPAAVGPTTMASGGGAPMA 212	
Db	126 PAPRHIFINLAKKAKGTSNTNSMSENVQHNHPINAGTEL-SATGNESGGGGGG--- 181	
Qy	213 DNNEGADGVGNASGNWH--CDSTWLG-----RVITSTRTWALPTYNNHLYKQIS--SAST 265	
Db	182 -GGRGAGGVGVSTGSGFNNQTEFYQLGEGLVIRITAHASRLIHLNMPEHETVKRIHVLNSES 240	
Qy	266 GAS-----NDNHVFGYSTPWGYFDNRFHCHFSPRDWQRLINNWWGRFPRKRLNFKLFNIQV 321	
Db	241 GVAGQMVDDAHTQMTVPWSLIDANAWGVWFPNADWQLISNNMTEINLVSFEQEIFNVVL 300	
Qy	322 KEVT--TNDGVTTIANNLSTVQVFSDBEYQLPYVLGSAHQGCLPPFPADVFMIPOYGY 378	
Db	301 KTITESATSPPTKTYNNDLTASLMVALDTNNTLPYTPAAPRSETLGFYPWLPKPTQYRY 360	
Qy	379 L-----TLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNPF-TFSYTPE 417	
Db	361 YLSCTRNLNPPYTGOSQOITDSIQTGLHSDIMFYTIENAVPIHLRLRTGDEFSTGIYHFD 420	
Qy	418 EVPFHSSYAHSSQLDR-----LMNPLI--DQYLYLNRNTQNSGSAQNKDLLFSRGS 467	
Db	421 TKPL--KLTHSWQTNRSGLPLPKLLTEPTTEGQHPGTLPAANTRKGYHOTMNNNSYTEAT 478	
Qy	468 PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMA 525	
Db	479 ----AIRP-----AQVGYNTPYMNFYSNGGPF-----LTPIVPTADT 512	
Qy	526 SHKDDKFFPM5GVMIFGKESAGASNTALDNVMTDEEEIKATNP-----VATERFGT 579	

Db 513 QYNDDE-----PNGAIRFTMGYQHGLTTS-----SQELERYTFNPQSKCGRAPKQFQNFQ 562

Qy 580 VA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVLQGPPIWAKIPHT 625

Db 563 QAPLNLENTNGTLLPSDPIGGKPNMHFMNTLNTYGPLTALNNTAPVFPNGQIWDKELDT 622

Qy 626 D--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPAEFSA-TKFASFITQYSTGQVSV 681

Db 623 DLKPRHL---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQOPRIITYSNFVWK 676

Qy 682 EIEWELQKENSQRNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728

Db 677 TLTFTAKMRSSNMWNPISQHTT-----TAENIGNYI-PTNIG 712

RESULT 5

COA1_PAVPN STANDARD; PRT; 729 AA.

AC P18546; P22964; Q89816;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Coat protein VP1 [Contains: Coat protein VP2].

OS Porcine parvovirus (strain NADL-2) (PPV).

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OX NCBI_TaxID=10797;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90085785; PubMed=2596019;

RA Vasudevacharya J., Basak S., Srinivas R.V., Compan R.W.;

RT "Nucleotide sequence analysis of the capsid genes and the right-hand

RT terminal palindrome of porcine parvovirus, strain NADL-2.";

RL Virology 173:368-377(1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91021005; PubMed=2219713;

RA Vasudevacharya J., Basak S., Srinivas R.V., Compan R.W.;

RT "The complete nucleotide sequence of an infectious clone of porcine

RT parvovirus, strain NADL-2.";

RL Virology 178:611-616(1990).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=90010964; PubMed=2794971;

RA Ranz A.I., Mancus J.J., Diaz-Aroca E., Casal J.I.;

RT "Porcine parvovirus: DNA sequence and genome organization.";

RL J. Gen. Virol. 70:2541-2553(1989).

CC -|- SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED

CC VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.

CC -|- MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT

CC OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.

CC -|- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M32787; AAA46917.1; -

DR EMBL; M32787; AAA46918.1; -

DR EMBL; M38367; AAA46919.1; -

DR EMBL; M38367; AAA46921.1; -

DR EMBL; D00623; BAA00502.1; -

DR PIR; B33302; VCPVPP.

DR PIR; B33743; VCPVNA.

DR HSSP; P30129; 4DPV.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

KW Coat protein; Glycoprotein.

FT CHAIN 1 729 COAT PROTEIN VP1.

FT CHAIN 151 729 COAT PROTEIN VP2.

FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 573 573 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 604 604 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT CARBOHYD 651 651 N-LINKED (GLCNAC. .) (POTENTIAL) .

FT DOMAIN 175 190 GLY-RICH.

FT CONFLICT 10 10 G -> GKGSEKGVVAYILQIIFYITG (IN REF. 3) .

FT CONFLICT 16 18 YKY -> TI (IN REF. 3) .

FT CONFLICT 56 56 Y -> T (IN REF. 3) .

FT CONFLICT 164 164 G -> A (IN REF. 3) .

FT CONFLICT 195 195 T -> S (IN REF. 3) .

FT CONFLICT 242 242 V -> S (IN REF. 3) .

FT CONFLICT 294 294 E -> A (IN REF. 3) .

FT CONFLICT 705 705 N -> K (IN REF. 3) .

FT CONFLICT 715 715 R -> K (IN REF. 3) .

SQ SEQUENCE 729 AA; 80946 MW; EF816E246C80DB42 CRC64;

Query Match 12.1%; Score 482; DB 1; Length 729;

Best Local Similarity 25.5%; Pred. No. 2.8e-22;

Matches 197; Conservative 111; Mismatches 310; Indels 154; Gaps 35;

Qy 43 RGLVLPQYKYLGPFGNLDKGEFVNAAADAALEHDKAYDQOLKAGDNPYLRYNHADAEPQE 102

Db 9 RGLTLPQYKYLGPFGNLDQGEPTNPSDAAAKEHDEAYDKYSGKNPYFYFSAADEKFIK 68

Qy 103 RLQEDTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGK- 161

Db 69 ETEHAKDYGGKIGHYFFRAKRAFAPKL---SETDSPTSQQPEVRRSPRGHPGSKPPGKR 125

Qy 162 -----TGQQPAKKRLNFGQTGDSSEVPDPQPLGEPATPAAGVPTTMASSGGGAPMA 212

Db 126 PAPRHIFINLAKKAKGTSTNSMSSENVEQHNPINAGTEL-SATGNESGGGGGGG--- 181

Qy 213 DNNEGADGVGNASG--NWHCDSTWLGD---RVITTSTRTWALPTYNNHLYKQIS--SAST 265

Db 182 -GGRGAGGVGVSTGTENNQTEFYLGEGLVRLTAHASRLHLNMPHEHETIKRIHVLNSES 240

Qy 266 GAS-----NDNHFGYSTPWGYDFENRPHCFSPRDWQRLNNWNGFRPKRLNFKLFNIQV 321

Db 241 GVAGQMVQDDAHTQMTVPWSLIDANAWGVWFNPADWQLISNNNTEINLVSEFEIENFVVL 300

Qy 322 KEVT---TNDGVTTIANNLSTVQVFSSEYQLPVLGSAHQGLPPPPADVFMIPQYGY 378

Db 301 KTITESATSPPTKIYNNDLTASLMVALDTNNTLPYTPAAPRSETLGFYPLWLPKPTQYRY 360

Qy 379 L-----TLNNGSQAVGRS-----SFYCLE-YFPQOMLRTGNMF-TFSYTPE 417

Db 361 YLSCIRNLNPPTYTGQSQITDSIQTLGHSIDIMFYTIENAVPIHLRTGDEFSTGIYHFD 420

Qy 418 EVPFSSVAHSQSLDR-----LMNPLI--DQYLYLNRNTQNSGSAQNKOLLFSRGS 467

Db 421 TKPL--KLTHSWQTNRSGLPLPKLLTEPTTEGDQHPGTLPAANTRKGYHQTNNSYTEAT 478

Qy 468 PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMA 525

Db 479 ----AIRP-----AQVGYNTPYMNFYSNGGPF-----LTPIVPTADT 512

Qy 526 SHKDEDEKFFPMGVMIFGKESAGASNTALD---NVMITDEEEK--ATNP-----VAT 574

Db 513 QYNDDEPN-----GAIRFTMDYQHGLTTSQELERYTFNPQSKCGRAPK 557

Qy 575 ERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVLQGPPIWA 620

Db 558 QQFNQQAPLNLENTNGTLLPSDPIGGKSNMHFMNTLNTYGPLTALNNTAPVFPNGQIWD 617

Qy 621 KIPHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPAEFSA-TKFASFITQYST 676

Db 618 KELDTDLKPRHL---VTAPFVCKNNPPGQLFVKIAP---NLTDDEFNADSPQOPRIITYSN 671

QY 677 GQVSVEIEWELQKENSQRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728
Db 672 FWWKGTLTFTAKMRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712

RESULT 6

COAL_PAVPK STANDARD; PRT; 729 AA.
AC P52501;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Porcine parvovirus (strain Kresse) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=73487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96183900; PubMed=8642680;
RA Bergeron J., Hebert B., Tijssen P.;
RT "Genome organization of the Kresse strain of porcine parvovirus:
RT identification of the allotropic determinant and comparison with
RT those of NADL-2 and field isolates.";
RL J. Virol. 70:2508-2515(1996).
CC -!- SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED
CC VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.
CC -!- MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT
CC OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U44978; AAC40230.1; -.
CC EMBL; U44978; AAC40231.1; -.
CC HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 729 COAT PROTEIN VP1.
FT CHAIN 151 729 COAT PROTEIN VP2.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 729 AA; 80835 MW; B6345BFA0568A1F6 CRC64;

Query Match 12.0%; Score 480; DB 1; Length 729;
Best Local Similarity 25.6%; Pred. No. 3.7e-22;
Matches 196; Conservative 112; Mismatches 315; Indels 144; Gaps 34;

QY 43 RGLVLPGYKYLPGFNGLDKGEVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFOE 102
Db 9 RGLTLPGYKYLPGNSLDQGEPTNPDAAKEHDEAYDKYKSNPYFYFSAADEKFIK 68
QY 103 RLQEDTSFGNLRGAVFOAKRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGK- 161
Db 69 ETEHAKDYGGKIGHYFFRAKRAFAPKL---SETDSPTTSQPEVRRSRPKHPGSKPPGKR 125
QY 162 -----TGQOPAKKRLNFGQTGDSSEVPDPQLGEPFATPAAVGPTTMASGGGAPMA 212
Db 126 PAPRHIFFINLAKKAKGTSTNTNSNSMSENVEQHNPINAGTEL-SATGNESGGGGGG--- 181

QY 213 DNNEGADGVGNASGNWH--CDSTWLGD---RVITTTSTRTWALPTYNNHLYKQIS--SAST 265
Db 182 -GGRGAGGVGVSTGSPNNQTEFYLGEGLVRIATAHARSLIHLNMPETHETRYKRIHVLNSES 240
QY 266 GAS----NDNHYFGYSTPWGYDFNRFHCHFSRPRDWRQLNNNWGFRPKRLNFKLPNIQV 321
Db 241 GVAGQMVQDDAHTQMTVPWSLIDANAWGVWFNPADQLISNNMTEINLVSFEQEIPNVVL 300
QY 322 KEVT--TNDGVTTIANLSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPQYGY 378
Db 301 KTITESATSPPTKIYVNDLTASLMVALDTNNTLPTYPAPRSETLGFYPWLPTKPTQYRY 360
QY 379 L-----TLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNMF-TESYTFE 417
Db 361 YLSCTRNLPPTYTGSSQIITDSIQTGLHSDIMFYTIENAVPIHLLRTGDEFSTGIYHFD 420
QY 418 EVPFHSSYAHSQSLDR-----LMNPLI--DQVLYLNRNTQNSGSAQNKDLLFSRGS 467
Db 421 TKPL--KLTHSWQTNRSGLGLPPKLLTEPTTEGDOHPGTLPAANTRKGYHOTINNSYTEAT 478
QY 468 PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSFTWTGASKYNLNGRESIINP--GTAMA 525
Db 479 ----AIRP-----AQGYNTPYMNFYSNGGPF-----LTPIVPTADT 512
QY 526 SHKDEDEKFFPMGVMIFGKESAGASNTALDNVMTDDEEIKATNP-----VATERFGT 579
Db 513 QYNDDE---PNGAIRFTMGYHGQLTTS-----SQELERYTFNPQSKCGRAPKQFNQ 562
QY 580 VA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVLQGPWAKIPHT 625
Db 563 QAPLNLENTNNGTLLPSDPIGGKPNMHFMNTLTYGTLTALNNTAPVFPNGQIWDKELDT 622
QY 626 D--GHFHPSPLMGGFGLK-NPPPIILIKNTVPVPANPPAEFSA-TKFASFITQYSTGQVSV 681
Db 623 DLKPRLLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQOPRIITYSNFWWKG 676
QY 682 EIEWELQKENSQRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 728
Db 677 TLTFTAKMRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712

RESULT 7

COAT_FPV STANDARD; PRT; 727 AA.
AC P04864; Q651112;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85265017; PubMed=2991581;
RA Carlson J., Rushlow K., Maxwell I., Maxwell F., Winston S., Hahn W.;
RT "Cloning and sequence of DNA encoding structural proteins of the
RT autonomous parvovirus feline panleukopenia virus.";
RL J. Virol. 55:574-587(1985).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M10824; AAA47161.1; -.
DR EMBL; M10824; AAA47162.1; -.
DR PIR; A03701; VCPV1F.
DR HSSP; P30129; 4DPV.

QY	324	VT---	TNDGVT	TIANNL	ISTVQV	FSDSEYQLPYVL	GSAHQCL-	-----PPFPA	-----	368
Db	291	VSESATQPPTKVVN	DLTASLMVALD	SNNMT	PFTPAAMR	SETLGFPY	KPFTIPTPWRYYP	350		
QY	369	--DVFMIPQGYLT	LNNGSQA	VGRS-	-----SPCYCLE-YFP	SQMLRTGNF-TFSY	414			
Db	351	QWDRTLIPSH-	-----TGTSCTPTNI	YHGTDPDDVQFY	TIENSVPVHL	LRTGDEFATGTF	404			
QY	415	TPEEVFFHSSYAHSQSL	DRLMN--PLIDQYL	LYLNRTQNQS	GSAQNKD	LLFSRGSPAGMS	472			
Db	405	FFDCKP--CRLTH	TWTNRALGLPP-	-----FLNSLPQ	SEGATNF	GDIGVQQDKRRGV	455			
QY	473	-----VQPKN-WLP	GPCYRQRVSK-	-----TKTDN	NNSNFTWTG	505				
Db	456	QMGNTRYITEATIMR	PAEVSAPYYSFEASTQG	PFKTPIAAGRGG	AQTENQAA--DG	512				
QY	506	ASKY---NLNGRESII	NPGT---AMASHKDEDEDKF	PFMSGVMIFGKS	SAGASNTALD-N	557				
Db	513	DPRYAFGRQHKGKTTT	TGETPERFTYIAHQDT-	-----GRYPAGDWIQ	NINFN	559				
QY	558	VMITDEEIKATNPVATER	FGTVAVN	FQSSSTDPATGDVHAMGALPGMWQDRD	VYLQGP	617				
Db	560	LPVTNDNVLLPTDP	IG---GKTGINY--TNIFNTYG	LTLANNVP-----PVYPNGQ	606					
QY	618	IWAKIPHDTGHFHP	SPLMGFGFLK-	-----NPPPQILIKNT	VPAN---PPAE	661				
Db	607	IWDKEPD	TD-----LKPR	LHVNAFPVCQNNCPGQLFVKVAPNL	TNEYDPDAS	653				
QY	662	FSATKFASFITQYSTG	VSVEIEWELQKENS	KRWNEPVQYTSNYAKSANVD	FTVDNNGLY	721				
Db	654	ANMSR----	IVTYSDFWWKGLVFKAKLRASH	TWNPIQOMSIN-----VDNQFNY	699					
QY	722	TEPRPIG	728							
Db	700	L-PNNIG	705							

```

RESULT 9
COAT_FPV19
ID COAT_FPV19 STANDARD; PRT; 727 AA.
AC P24840;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein (FPV)].
OS Feline panleukopenia virus (strain 193) (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=193/70;
RX MEDLINE=91073139; PubMed=2174965;
RA Martyn J.C., Davidson B.E., Studdert M.J.;
RT "Nucleotide sequence of feline panleukopenia virus: comparison with
RT canine parvovirus identifies host-specific differences.";
RL J. Gen. Virol. 71:2747-2753(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CU-4;
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine
RT parvovirus and feline panleukopenia virus using infectious plasmid
RT clones.";
RL Virology 183:195-205(1991).
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

[illegible]

```
Db 611 QIWDKEFTD-----LKPRHLVNAFFVQCNCPGQLFVKVAPNLNTEYDPDA 657
Qy 661 EFSATKFPASFTQYSTQVSVEIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNGL 720
Db 658 SANMSR-----IVTYSDFWKKGLVFKAKLRASHWTWNPIQOMSIN-----VDNQFN 703
Qy 721 YTEPRPIG 728
Db 704 YV-PNNIG 710

RESULT 10
COAT_PAVCB
ID COAT_PAVCB STANDARD; PRT; 727 AA.
AC Q11213;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Canine parvovirus (strain B) (CPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=59284;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine parvovirus and feline panleukopenia virus using infectious plasmid clones.";
RL Virology 183:195-205(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38245; AAB02799.1; -.
CC DR EMBL; M38245; AAB02800.1; -.
CC DR HSSP; P30129; 4DPV.
CC DR InterPro; IPR001403; Parvo coat.
CC DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 727 COAT PROTEIN VP1.
FT CHAIN 144 727 COAT PROTEIN VP2.
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 660 660 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 727 AA; 80342 MW; 115F3E1A79098EBE CRC64;

Query Match 10.8%; Score 430.5; DB 1; Length 727;
Best Local Similarity 23.2%; Pred. No. 4e-19;
Matches 184; Conservative 112; Mismatches 297; Indels 201; Gaps 32;

Qy 43 RGLVLPGYKYLGPFGNLDKGEFVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE 102
Db 10 RGLVPPGYKYLGPFGNLDQGEPTNPSDAAAKEHDEAYALRSKGNPYLYFSPADQRFID 69
Qy 103 RLQEDTSFGGNLGRAVQAKKRVLEPLGLVEEAKTAPGKRPVEQSPQEPDSSSGIGKT 162
Db 70 QTKDAKDWGKGIGHYFFRAKKAIAPLVLTDPDHPSPSRPTK-PTKRSKPPPHIFINLAK- 127
Qy 163 GQPAKKRLNFGQTGDSSEVP-----DPQPLGEPPA--TPAAVGPPTTMASGGGAPMADNNE 216
```

```
Db 128 -----KKKAGAGQVKRDNLAPMSDGAQVQPDGGQPAVNERATSGNSGGG-----GG 176
Qy 217 GADGVGNASGNWH-----CDSTWLGDRIITTTSTRTWALPTYNHLYKQI-----SS 262
Db 177 GSGVGISTGTFFNQTEFEKLENGWV--EITANSSRLVHLNMPESYRVRVNNMDKTA 234
Qy 263 ASTGASNDNHYFGYSTPWGYFDNRFCHFSRDPDQRLNNWGFPRKRLNFKLFNIQVK 322
Db 235 VNGNMALDDIHAQIVTPWSLVLDANAWGVFNPGDQQLIVNTMSELHLVSFEQEIFNVVLK 294
Qy 323 EVT---TNDGVTTIANLSTVQVPSDSEYQLPYVLGSAHQGCL-----PPFPA----- 368
Db 295 TVSESATQPPTKVYNNDLTASLMVALDSNNTMPFTPAAMRSETLGFYPWKPTIPTPWRY 354
Qy 369 ---DVFMIPOGYGLTLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNF-TFS 413
Db 355 FQWDRTLIPSH-----TGfSGTPTNIYHGTDPPDDVQFYTIENSVPVHLLRTGDEFATGT 408
Qy 414 YTFEEVFPFHSSYAHSQSLDRIMN--PLIDQYLYLNRTONQSGSAQNKOLLFSRGSPAGM 471
Db 409 FFFDCKP--CRLTHTWTQTNRALGLPP-----FLNSLPQSEGATNFGDI-----GV 451
Qy 472 SVQPKNWLPGCYRQQRVSKTKTDNNNSNF-----TWTGASKYNLNGRESIINP-- 520
Db 452 -----QDDKRRGVTTQMGNTNYITEATIMRPAEVGVSAPYYSFEASTQGPFK 497
Qy 521 ---GTAMASHKDDKEDKFFPMSGVMI FGKESAGASNTALD----- 556
Db 498 TPIAAGRGAQTDENQAADGNRYAFGRHQGKTTTGTETPERFTYIAHQDTGRYPEGDW 557
Qy 557 -----NVMITDEEIKATNPVATERFGTVAVNFQSSSTDATGDVHAMGALPGMVWQDR 610
Db 558 IQNINENLPVTNDNVLLPTDPG----GKTGINY--TNIFNTYGLTALNNVP----- 604
Qy 611 DVYLGPIWAKIPHDTGHFHPSPLMGGFGLK-----NPPQILIKNTPPVAN 657
Db 605 PVYPNGQIWDKEFTD-----LKPRHLVNAFFVQCNCPGQLFVKVAPNLTN 651
Qy 658 ---PPAEFSATKFPASFTQYSTQVSVEIEWELQKNSKRWNPVEQYTSNYAKSANVDFT 714
Db 652 EYDPDASANMSR---IVTYSDFWKKGLVFKAKLRASHWTWNPIQOMSIN----- 697
Qy 715 VDNNGLYTEPRPIG 728
Db 698 VDNQFNIV-PSNIG 710

RESULT 11
COAT_MUMIM
ID COAT_MUMIM STANDARD; PRT; 718 AA.
AC P07302; Q9WMH2; Q9WMH3;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Murine minute virus (strain MVMi) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain.";
RT J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice.";
```



```
Db      648 SRIVTYGTFWKGKLTMRACKLRA-----NTTNVPVYQ-----VSVEDNG 686
      ::  ::  |  :  |:::  :  |||  |  :  :::  :  :::  :  :::  :
RESULT 12
COAT_PAVCN
ID  COAT_PAVCN  STANDARD;  PRT;  748 AA.
AC  P12930; Q84387;
DT  01-OCT-1989 (Rel. 12, Created)
DT  01-OCT-1989 (Rel. 12, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Coat protein VP1 [Contains: Coat protein VP2].
OS  Canine parvovirus (strain N) (CPV).
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10791;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88062992; PubMed=2824850;
RA  Reed A.P., Jones E.V., Miller T.J.;
RT  "Nucleotide sequence and genome organization of canine parvovirus.";
RL  J. Virol. 62:266-276(1988).
CC  -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC  OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC  -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M19296; AAA67460.1; -
DR  EMBL; M19296; AAA67461.1; -
DR  PIR; B29962; VCPVCP.
DR  HSSP; P30129; 4DPV.
DR  InterPro; IPR001403; Parvo_coat.
DR  Pfam; PF00740; Parvo_coat; 1.
KW  Coat protein; Glycoprotein.
FT  CHAIN 1 748 COAT PROTEIN VP1.
FT  CHAIN 165 748 COAT PROTEIN VP2.
FT  CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  DOMAIN 186 203 GLY-RICH.
FT  CONFLICT 27 27 M -> XLSM (IN REF. 1; AAA67460).
SQ  SEQUENCE 748 AA; 82715 MW; 13E487C6862C73E6 CRC64;
Query Match 10.6%; Score 421.5; DB 1; Length 748;
Best Local Similarity 23.0%; Pred. No. 1.5e-18;
Matches 184; Conservative 114; Mismatches 287; Indels 215; Gaps 34;
Qy 44 GLVLPGYKYLGPENGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEPQER 103
Db 32 GLVPPGYKYLGPNSLDQGEPTNPSDAAAKEHDEAYAYLRSGKNPYLYFSPADQRFIDQ 91
Qy 104 LQEDTSFGNGLGRAVFOAKRVLRLGLVEEGAKTAPGKRKPVQSPQEPDSSSIGIKTG 163
Db 92 TKDAKDWGGKIGHYFFRAKKAIAPLVLTTPDHPSTSRPTK-PTKRSKPPPHIFINLAK-- 148
Qy 164 QQPAKKRLNFGQTGDSVSP---DPQPLGEPPA--TPAAVGTPTWASGGGAPMADNNEG 217
Db 149 ----KKKAGAGQVKRDNLAPMSDGAQVDPGGQPAVRNERATSGNGSGGGG-----GGG 198
Qy 218 ADGVGNASGNWH-----CDSTWLGDRVITTTSTRTWALPTVNNHLYKQI-----SSA 263
Db 199 SGGVGIGISTGFNNQTEFKFLENGWV--EITANSSRLVHLNMPSENYRRVVNNMDKTAV 256
```

```
Qy 264 STGASNDNHYFGYSTPMGYFDENRFHCHFSPRDWQRLINNNGFRPKRLNFKLFNIQVKE 323
Db      :  :  :  |  :  |:::  :  |||  |  :  :::  :  :::  :  :::  :
Db 257 NGNMALDDIHAQIVTPWSLVDANAGWGFNPGDWQLIVNTMSELHLVSFEQEIFNVVLKT 316
      :  :  :  |  :  |:::  :  |||  |  :  :::  :  :::  :  :::  :
Qy 324 VT---TNDGVTTIANLSTVQVFSDSYQLPYVLGSAHQCL-----PPFPA----- 368
Db      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 317 VSESATQPTKPVYNNDLTASLMVALDSNNTMPFTPAAMRSETLGFYPWKPTIPTPWRYF 376
      :  :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Qy 369 --DVFMIPQYGYLTNNGSQAVGRS-----SFYCLE-YFYSQMLRTGNMF-TFSY 414
Db      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 377 QWDRTLIPSH-----TGTSGTPTNIYHGTDPDDVQFYTIENSVPVHLLRTGDEFATGF 430
      :  :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Qy 415 TFEVVPFHSSYAHQSLSLRLMN--PLIDQYLYLNRNTQNSGSAQNKDLLFSRGSAGMS 472
Db      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 431 FFDCKP--CRLTHTWQTNRALGLPP-----FLNSLPQSEGATNFGDI-----GV- 472
      :  :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Qy 473 VQKXWLPQPCYRQORVSKTKTDNNNSNF-----TWTGASKYNLNGRESIINP-GT 522
Db      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 473 -----QODKRRGVTTQMGNTNYITEATIMRPAEVGYSAFYSPPEASTQGPFKT 519
      :  :  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Qy 523 AMASHK-----DDEKFFPMGVMIFGKESAGASNTALD----- 556
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Db 520 PIAAGRGGAQTYENQAADGDPY-----AFGRHQGQKTTTGTGETPERFTYIAHQDTGR 572
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Qy 557 -----NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPTDPTGDVHAMGALPG 604
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Db 573 YPEGDWIQINFNLPVTNDNVLLPTDPIG---GKTGINY--TNIFNTYGPLTALNNVP- 625
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Qy 605 MVWQDRDVYLQGIPIWAKIPHDTGHPHPSPLMGFGGLK-----NPPQILIKN 651
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Db 626 -----PVYPNGQIWDKEFTD-----LKPRLHVNAFPVCQNNCPGQLFVKV 666
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Qy 652 TPVPAN---PRAEFSATKFASFITQYSTQVSVEIEWELQKENSXKWNPEVQYTSNYAKS 708
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Db 667 APNLTNEYDPDASANMSR---IVTYSDFWVKGLVFKAKLRASHTWNPIQQMSIN---- 718
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Qy 709 ANVDFTVDNNGLYTEPRPIG 728
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
Db 719 -----VDNQFNVY-PSNIG 731
      :  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |  :  :  |
RESULT 13
COAT_PAVCN
ID  COAT_PAVCN  STANDARD;  PRT;  737 AA.
AC  P17455;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Coat protein VP1 [Contains: Coat protein VP2].
OS  Canine parvovirus (strain CPV-D Cornell 320) (CPV).
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX  NCBI_TaxID=10790;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89020796; PubMed=3176341;
RA  Parrish C.R., Aquadro C.F., Carmichael L.E.;
RT  "Canine host range and a specific epitope map along with variant
RT  sequences in the capsid protein gene of canine parvovirus and related
RT  feline, mink, and raccoon parvoviruses.";
RL  Virology 166:293-307(1988).
CC  -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC  OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC  -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M23255; AAA47158.1; -
```


QY 327 -----NDGVTTIANLSTVQVFSQSEYQLPYVLGSAHQGCL-----PPFPAD---VFMI 373
Db 291 QGAGQDAIKVYNNDLTACMMVALDSNNILPYTPAAQTSETLGFYPWKPTAPAPRYFFEM 350
QY 374 PQYGLYTLNNGSQ-----AVG-----RSSFYCLE-YFPQSMLRTGNNPTF-SYTFEEVP 420
Db 351 PRQLSVTSSNSAEGTQITDTIGEPQALNSQFFTIENLTPLTLRTGDEFTGTIFYFNTDP 410
QY 421 FHSSYA-----HSQSLDRLMN-PLIDQYLYLNRNQSGSAQNKLLF----- 463
Db 411 LKLTHTWQTNRHLACLQGITDLPTSDTATASLTANGDRFGSTQTQNVNVTVEALRTRPAQ 470
QY 464 -----SRGSPAGMSVQPKWLPGPCYRQORVSKTKTDNNNSNFTWTGASKYNL 511
Db 471 IGFMQPHDNFEANRGGPFKVPVVP-----LDITAGEDHDAN---GAIRFNY 513
QY 512 NGR--ESIINPGTA-----MASHKDDDEDKFFPMGVMIFGKESAGASNTALDNVM 559
Db 514 GKQHGEDWAKQGAAPERYTWDIDAIDSAAGRDTCRCFV-----QSAPISIPPNQOI 563
QY 560 ITDEEEIKA-TNPVATEREGTVAVNFQSSSTDPTATGDVHAMGALPGMVWQDRDYYLQGP 618
Db 564 LOREDAIAGRTNMHYTNVFNYSGLSAPHPDP-----IYPNGOI 603
QY 619 WAK---IPHTDGHFHPSPMLGCGFGLKNPPPPQILIKNTPVPANPPAEP--SATKFASFITQ 673
Db 604 WDKELDLHKPRLHVTAPFV---CKNPPGQLFVHLGP---NLTDQFDPNSTVSRIVT- 656
QY 674 YSTGQVSVEIEWELQKENSKRWNPEVQYTSN 704
Db 657 YSTFYWKGLKFKAKLRPNLTWNPVYQATTD 687

RESULT 15
COAT_PAVC7
ID COAT_PAVC7 STANDARD; PRT; 722 AA.
AC P04863;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Canine parvovirus (strain 780929) (CPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10789;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185696; PubMed=3989914;
RA Rhode S.L. III;
RT "Nucleotide sequence of the coat protein gene of canine parvovirus."; J. Virol. 54:630-633 (1985).
RL [2]
RN REVISIONS.
RA Rhode S.L. III;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 228-473 FROM N.A.
RX MEDLINE=86098657; PubMed=3942033;
RA Parrish C.R., Carmichael L.E.;
RT "Characterization and recombination mapping of an antigenic and host range mutation of canine parvovirus."; Virology 148:121-132 (1986).
RL
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M10989; AAA47149.1; -.
DR EMBL; M11871; AAA47147.1; -.
DR PIR; A03702; VCPVCN.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 722
FT CHAIN 139 722
FT CARBOHYD 163 163
FT CARBOHYD 185 185
FT CARBOHYD 202 202
FT CARBOHYD 318 318
FT CARBOHYD 581 581
FT CARBOHYD 643 643
FT CARBOHYD 655 655
FT DOMAIN 160 181
FT CONFLICT 409 409
FT CONFLICT 438 439
FT CONFLICT 447 447
FT CONFLICT 451 451
FT CONFLICT 454 454
SQ SEQUENCE 722 AA; 79826 MW; BBB37380B4371429 CRC64;
Query Match 9.7%; Score 386.5; DB 1; Length 722;
Best Local Similarity 22.3%; Pred. No. 2e-16;
Matches 176; Conservative 111; Mismatches 300; Indels 201; Gaps 32;
QY 49 GYKYLGFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPLYRNHADAERLQEDT 108
Db 11 GYKYLGPGLDQGEPTNPSDAAAKEHDEAYAYLRSGKNPLYFSPADQRFIDQTKDAK 70
QY 109 SFGGNLGRAVFQAKKRVLEPLGLVEEAKTAPGKRVEQSPQEPDSSSGIGKTQQPAK 168
Db 71 DWGGKIGHYFFRAKKAIAPLVLTDPDHPSTSRPTK-PTKRSKPPPHIFINLAK-----K 123
QY 169 KRLNFGQTGDSSEVP---DPQLGEPPEA--TPAAVGPTTMSGGGAPMADNNEGADGVG 222
Db 124 KKAGAGQVKRDNLAPMSDGAVPDGGQPAVRNERATSGNGSGGG-----GGSGGVG 177
QY 223 NASGNWH-----CDSTWLGDVRVITSTRTWALPTYNNHLYKQI-----SSASTGAS 268
Db 178 ISTGTENNQTEFKLENGWV--EITANSSRLVHLNMPSESEKDRRVVNNMDKTAVNGNMA 235
QY 269 NDNHYFGYSTPWGYFDNRFHCHESPRDWQRLLNNWGRPKRLNFKLNIQVKEVT--- 325
Db 236 LDDIHAQIVTPWLSLVDANAWGVWFNPGDWQLIVNTMSELHLVSFEQEIFNVVLKTVSESA 295
QY 326 TNDGVTTIANLSTVQVFSQSEYQLPYVLGSAHQGCL-----PPPEA-----DVF 371
Db 296 TQPPTKVYNNDLTASLMVALDSNNTMPPTPAAMRSETLGFYPWKPTIPTPWRYYPQWDR 355
QY 372 MIPQGYLTLNNGSQAVGRS-----SFYCLE-YFPQSMLRTGNNF-TFSYTPPEEV 419
Db 356 LIPSH-----TGTSGTPTNIYHGTDPPDVQFYTIENSVPVHLLRTGDEFAFGTFFDCK 409
QY 420 PFHSSYAHQSGLDRLMN--PLIDQYLYLNRNQSGSAQNKOLLFSRGSPPAGMSVQPKN 477
Db 410 P--CRLTHTWQTNRALGLPP-----FLNSLPQSEGATNFGDI---GVP----- 447
QY 478 WLPGPCYRQQRVSKTKTDNNNSNF-----TWTGASKYNLNGRESINP-----GTA 523
Db 448 -----QDKKRGVTQMGTNTYITEATIMRPAEVGYSAPYYSFEASTQGFKTPIAAG 498
QY 524 MASHKDDDEDKFFPMGVMIFGKESAGASNTALD----- 556
Db 499 RGAQTDENQAADGNPRYAFGRQHGKKTITTTGETPERFTYIAHQDTGRYPEGDWIQNINF 558
QY 557 NMVITDEEEIKATNPVATERFGTVAVNFQSSSTDPTATGDVHAMGALPGMVWQDRDYYLQ 616
Db 559 NLPVTNDNVLLPIDDIG---GKTGINY--TNIFNTYGLPLTALNNVP-----PVYPNG 605
QY 617 PIWAKIPHTDGHFHPSPMLMGFGFLK-----NPPPOILIKNTPVPAN---PPA 660

Db	606	QIWDKEFTD-----LKPRLHVNAPFVCQNNCPGQLFVKLAPNLTNEYDPDA	652
QY	661	EFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVOYTSNYAKSANVDFTTVDNNGL	720
Db	653	SANMSR---IVTYSDFWKGKLVFKAKLRASHTWNPIQQMSIN-----VDNQFN	698
QY	721	YTERPIG	728
Db	699	YV-PSNIG	705

Search completed: January 21, 2004, 16:02:10
Job time : 14.6014 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:01:27 ; Search time 16.9331 Seconds
(without alignments)
1839.046 Million cell updates/sec

Title: US-09-807-802A-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3402.5	85.3	735	4 US-09-321-589-1	Sequence 1, Appli
2	2486.5	62.3	734	4 US-09-532-594B-4	Sequence 4, Appli
3	1830.5	45.9	598	4 US-09-532-594B-16	Sequence 16, Appli
4	1690.5	42.4	544	4 US-09-532-594B-18	Sequence 18, Appli
5	1410	35.3	756	4 US-09-438-268-4	Sequence 4, Appli
6	479.5	12.0	543	3 US-08-856-841-22	Sequence 22, Appli
7	430.5	10.8	500	3 US-08-856-841-16	Sequence 16, Appli
8	430.5	10.8	501	3 US-08-856-841-18	Sequence 18, Appli
9	429.5	10.8	486	3 US-08-856-841-19	Sequence 19, Appli
10	326	8.2	415	3 US-08-856-841-20	Sequence 20, Appli
11	318	8.0	395	3 US-08-856-841-13	Sequence 13, Appli
12	310	7.8	264	3 US-08-856-841-14	Sequence 14, Appli
13	303	7.6	398	3 US-08-856-841-21	Sequence 21, Appli
14	284	7.1	387	3 US-08-856-841-17	Sequence 17, Appli
15	216.5	5.4	579	6 5223424-13	Patent No. 5223424
16	192	4.8	584	3 US-09-022-949-2	Sequence 2, Appli
17	142	3.6	210	3 US-08-856-841-9	Sequence 9, Appli
18	142	3.6	227	3 US-08-856-841-15	Sequence 15, Appli
19	142	3.6	250	3 US-08-856-841-12	Sequence 12, Appli
20	124	3.1	3060	2 US-08-487-826B-14	Sequence 14, Appli
21	121	3.0	2736	4 US-09-252-991A-30227	Sequence 30227, A
22	119.5	3.0	1394	3 US-08-296-791-2	Sequence 2, Appli
23	119.5	3.0	1394	5 PCT-US95-10661A-2	Sequence 2, Appli
24	119	3.0	1095	4 US-09-107-532A-3855	Sequence 3855, Ap
25	114.5	2.9	434	2 US-08-710-249-4	Sequence 4, Appli
26	114.5	2.9	434	4 US-09-220-157A-4	Sequence 4, Appli
27	113.5	2.8	655	1 US-08-469-202-27	Sequence 27, Appli

28	113.5	2.8	655	2 US-08-484-434C-34	Sequence 34, Appli
29	113.5	2.8	655	4 US-09-384-361-34	Sequence 34, Appli
30	112.5	2.8	1096	4 US-09-415-946-3	Sequence 3, Appli
31	111.5	2.8	1245	4 US-09-252-991A-30935	Sequence 30935, A
32	110.5	2.8	624	3 US-08-947-965-78	Sequence 78, Appli
33	110.5	2.8	655	1 US-08-469-202-28	Sequence 28, Appli
34	110.5	2.8	655	2 US-08-484-434C-35	Sequence 35, Appli
35	110.5	2.8	655	4 US-09-384-361-35	Sequence 35, Appli
36	110.5	2.8	1751	3 US-09-136-574A-44	Sequence 44, Appli
37	109.5	2.7	1651	3 US-09-540-245A-18	Sequence 18, Appli
38	109.5	2.7	1848	3 US-08-296-791-6	Sequence 6, Appli
39	109.5	2.7	1848	5 PCT-US95-10661A-6	Sequence 6, Appli
40	109	2.7	1186	1 US-08-485-568A-4	Sequence 4, Appli
41	109	2.7	1186	1 US-08-357-698-6	Sequence 6, Appli
42	109	2.7	1186	2 US-08-590-554A-4	Sequence 4, Appli
43	109	2.7	1186	2 US-09-184-223-4	Sequence 4, Appli
44	109	2.7	1186	5 PCT-US93-12682-6	Sequence 6, Appli
45	109	2.7	1545	3 US-08-296-791-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.3%; Score 3402.5; DB 4; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.5e-289;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

Qy	1	MAADGYLPDWLEDNLSEGR	IREWDLKPGAPKPKANQKQDDGRGLVLP	PGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLETLSEGR	IQWVKLPGPPPKPAERHKDDSRGLVLP	PGYKYLGPFGNGLD	60
Qy	61	KGEPVNAADAAALEHDKAYDQQLKAGDNP	LYRYNHADA	AEFQERLQEDTSFGGNLGRAVFQ	120
Db	61	KGEPVNEADAAALEHDKAYDRQLDSDGNP	LYKYNHADA	AEFQERLKEDTSFGGNLGRAVFQ	120
Qy	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSP	QEPDSSSGIGTKGQQPAKKRLNFGQTGDSE	180	
Db	121	AKKRVLEPLGLVEEPVKTAPGKKRPVEHSP	VEPDPSSSGTGKAGQQPARKRLNFGQTGDAD	180	
Qy	181	SVDPDQPLGEPPTAAAVGPTTMA	SGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240	
Db	181	SVDPDQPLGQPPAAPSGLTNTMATG	SAPMADNNEGADGVGNSSGNWHCDSTWMDRVI	240	
Qy	241	TTSTRTWALPTYNHLYKQISSASTGASND	NHFGYSTPWGYFDFNRFHCHFSRQRL	300	
Db	241	TTSTRTWALPTYNHLYKQISSQS-GASND	NHFGYSTPWGYFDFNRFHCHFSRQRL	299	
Qy	301	INNNWGFRPKRLNFKLFNIQVKEVTNDG	VTTIANNL	TSVQVFS	360
Db	300	INNNWGFRPKRLNFKLFNIQVKEVTNDG	VTTIANNL	TSVQVFTDSEYQLPYVLGSAHQ	359
Qy	361	GCLPPFPADVFMIQYGYLTLNNGSQAVGR	SSFCLEYFPPSQMLRTGNNFTSYTFEEVP	420	
Db	360	GCLPPFPADVFVQYGYLTLNNGSQAVGR	SSFCLEYFPPSQMLRTGNNFTSYTFEDVP	419	

QY 421 FHSSYAHQSQSLDRLMNPLIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKWLP 480
Dd 420 FHSSYAHQSQSLDRLMNPLIDQYLYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKWLP 479
QY 481 GPCYRQORVSKTKTDNNNSFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGV 540
Dd 480 GPCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFFPMSGV 539
QY 541 MIFGKESAGASNTALDNVMTDDEEIKATNPVATERFCTVAVNFQSSSTDPATGDVHAMG 600
Dd 540 LIFGKQSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQGRNRQAATADVNTQG 599
QY 601 ALPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQIILIKNTPVPANPPA 660
Dd 600 VLPGMVWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKHPPPPQIILIKNTPVPANPST 659
QY 661 EFSATKFASFITQYSTGQSVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 720
Dd 660 TFSAAKFASFITQYSTGQSVSVEIEWELQKENSKRWNPEIQYTSNYNKS VNVDFTVDTNGV 719
QY 721 YTEPRPIGTRYLTRPL 736
Dd 720 YSEPRPIGTRYLTRNL 735

RESULT 2
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chordini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP1
US-09-532-594B-4

Query Match 62.3%; Score 2486.5; DB 4; Length 734;
Best Local Similarity 63.5%; Pred. No. 3.2e-209;
Matches 475; Conservative 81; Mismatches 161; Indels 31; Gaps 10;
QY 4 DGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPYKYLGPFGNGLDKGE 63
Dd 3 DGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNARGLVLPYKYLGPFGNGLDKGE 62
QY 64 PVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFOAKK 123
Dd 63 PVNAADAAALEHDKAYDQOLKAGDNPYLYKNHADADEFQORLOQDTSFGGNLGRAVFOAKK 122
QY 124 RVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKGTGQOPAKKRLNF-GQTGSESV 182
Dd 123 RVLEPLGLVEQAGETAPGKKRPLIESPQQPDSSTGIGKKGKQPAKKLVFEDETGAGDGP 182
QY 183 PDPQPLGEPPTAAVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTT 242
Dd 183 PEGSTSG-----AMSDDSEMRAAGAAVEGGQADGVGNASGDWHCDSTWSEGHVTTT 236
QY 243 STRTVALPTYNHLYKQISSASTGASNDNHFGYSTPWPYGFDFNRFCHFSRPRDWORLIN 302

Dd 237 STRTVALPTYNHLYKRLGE-----SLQSNTRYNGFSTPWPYGFDFNRFCHFSRPRDWORLIN 292
QY 303 NNWGFPRKRLNFKLFIQVKEVTNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQGC 362
Dd 293 NNWGMRPKAMRVKIFNIQVKEVTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGS 352
QY 363 LPFPADVEMIPQYGY--LTLNNGSQAVGRSSFCYCLEYFPQMLRTGNFTFSYTFEEV 419
Dd 353 LPFPNDVFMVPYGYCGLVTGNTSQQTDNRNFCYCLEYFPQMLRTGNFTFSYTFEY 412
QY 420 PFHSSYAHQSQSLDRLMNPLIDQYLYLNRTQNTQ-----QSGSAQNKDLLFSRGSFAGMSVQ 475
Dd 413 PFHSMYAHQSQSLDRLMNPLIDQYLGWGLQSTTTGTTLNAGTATTN---FTKL RPTNFSNFK 469
QY 476 KNWLPGPCYRQORVSKTKTDNNNSNFTWTGAS---KY-----NLNGRESINPGTAMASHK 528
Dd 470 KNWLPGPSIKQQGFSKTA--NQNYKIPATGSDSLIKYETHSTLDGRWSALTGPPPMATAG 527
QY 529 DDEDKFFPMSGVMIIFGKESAGASNTALDNVMTDDEEIKATNPVATERFCTVAVNFQSSS 588
Dd 528 PADSK-FSNSQLIFAGPKQNGNTATVPGTLIFTSEELAAATNATDMDMGNLPGGDQSN 586
QY 589 TDPATGDVHAMGALPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQIL 648
Dd 587 NLPTVDRLTALGAVPGMWQNRDIYQGPPIWAKIPHTDGHFHPSPMLIGFGLKHPPPPQIF 646
QY 649 IKNTVPANPPAFSATKFASFITQYSTGQSVSVEIEWELQKENSKRWNPEVQYTSNYAKS 708
Dd 647 IKNTVPANPATTFSSPTVNSFITQYSTGQSVSVQIDWEIQERSKRWNPEVQYTSNYGQQ 706
QY 709 ANVDFTVDNNGLYTEPRPIGTRYLTRPL 736
Dd 707 NSLLWAPDAAGKYTEPRAIGTRYLTRHL 734

RESULT 3
US-09-532-594B-16
; Sequence 16, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chordini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16

Query Match 45.9%; Score 1830.5; DB 4; Length 598;
Best Local Similarity 57.8%; Pred. No. 7.9e-152;
Matches 355; Conservative 74; Mismatches 154; Indels 31; Gaps 10;
QY 138 TAPGKKRPVEQSPQEPDSSSGIGKGTGQOPAKKRLNF-GQTGDSSEVPDPQPLGEPPTPA 196
Dd 1 TAPGKKRPLIESPQQPDSSTGIGKKGKQPAKKLVFEDETGAGDGPPEGSTSG-----A 54
QY 197 AVGPPTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTRTVALPTYNH 256
Dd 55 MSDDSEMRAAGAAVEGGQADGVGNASGDWHCDSTWSEGHVTTTSTRTRTVALPTYNH 114

QY	257	YKQISSASTGASNDNHFYGYSTPWGYDFDNRFCHCHFSPRDWORLNNWGWFRPKRLNFKL	316
Db	115	YKRLGE---SLQSNNTYNGFSTPWGYDFDNRFCHCHFSPRDWORLNNWGWMRPKAMRVKI	170
QY	317	FNIQVKEVTNDGVTTIANNLSTSTVQVFSDESEYQLPYVLGSAHQCLPPFPADVFMIPOY	376
Db	171	FNIQVKEVTTSNGETTAVANLSTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPOY	230
QY	377	GY---LTLNNGSQAAGRSSFYCLEYFPFSQMLRTGNFTFSYTFEEVPHSSHVAHSQSLLDR	433
Db	231	GYCGLVTGNTSQQOTDRNAFYCLEYFPFSQMLRTGNFETIYSFEKVPFHSMYAHSQSLLDR	290
QY	434	LMNPLIDQLYLYLNRTON---QSGSAQNKDLLFSRGSAPAGMSVQPKXWLPGPCYRQORV	489
Db	291	LMNPLIDQYLWGLQSTTGTTLNAGTATTN---FTKLRPTNFSNFKXWLPGPSIKQQGF	347
QY	490	SKTKTDNNNSFTWTGAS---KY---NLNGRESIINPGTAMASHKDDDEDKFFPMMSGVMI	542
Db	348	SKTA--NQYKIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSK-FSNSQLIF	404
QY	543	FGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGAL	602
Db	405	AGPKQNGNTATVPGTLIFTSEEELAAATNATDMDWGNLPGGDQSNLSLFTVDRRLTALGAV	464
QY	603	PGMVWQDRDVYLOGPWIWAKIPHDTGHFHPSPLMGGFGLKONPPQILIKNTPVPANPPAEF	662
Db	465	PGMVWQNRDIYYQGPWIWAKIPHDTGHFHPSPLIIGFGLKHPPPPQIFIKNTPVPANPATTF	524
QY	663	SATKPFASFITQYSTGOVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYT	722
Db	525	SSTPVSFITQYSTGOVSQVQIDWEIQKERSKRWNPEVQFTSNYQQNSLLWAPDAAGKYT	584
QY	723	EPRPIGTRYLTRPL	736
Db	585	EPRAIGTRYLTHHL	598

```

RESULT 4
US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18

```

[illegible]

```

RESULT 5
US-09-438-268-4
; Sequence 4, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Virus
US-09-438-268-4

```

		Query Match	35.3%;	Score 1410;	DB 4;	Length 756;
		Best Local Similarity	42.7%;	Pred. No. 8.8e-115;		
		Matches 312; Conservative	95;	Mismatches 269;	Indels 54;	Gaps 16
Qy	1	MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKDDGRGLVLPGYKYLGPFNGLD	60			
		:				
Dd	1	MAADGYLPDWLEDTLSEGIRQWVKLKPGPPPPPKPAERHKDDSRGLVLPGYKYLGPFNGLD	60			
		:				
Qy	61	KGEPVNADAAALEHDKAYDQQLKAGDNPYLRYNHADAEEQERLQEDTSFGGNLGRAVFQ	120			
		:				
Dd	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEEQERLQEDTSFGGNLGRAVFQ	120			
		:				

Db 259 MWSEGATFSANSVTCTFSRQFLIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGY 318
Qy 277 STPWGYFDENRPHCHFSRQDLINNNWGFPRKRLNFKLFNIQVKEVT--TNDGVTTIA 334
Db 319 STPWRYLDFNALNLFSPLEFQHLIENYGSIAPDALTVTIASEIAVKDVTDKTGGV-QVT 377
Qy 335 NNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTLNN-GSQAVG---- 389
Db 378 DSTTGRCLMLVDHEYKYPYVLGGQDTLAPELPIWVYFPPQYAYLTVGDVNTQGISGDSK 437
Qy 390 -----RSSFYCLEYFPQMLRTGNFTFSYTFEEVPPHSSYAHSQSLDRLMNPL 438
Db 438 KLAESAFYVLEHSSFQLLGTGTASMSYKFPVPVPPENLEGCSQH FYEMYNPL 491

RESULT 8
US-08-856-841-18
; Sequence 18, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: N/A
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
; IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
; POSITION IN GENOME: N/A

FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; AUTHORS: WIDDOWS, D.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET
; VOLUME: I
; ISSUE:
; PAGES: 72 - 73
; DATE: 1975
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 18:
US-08-856-841-18

Query Match 10.8%; Score 430.5; DB 3; Length 501;
Best Local Similarity 29.5%; Pred. No. 3.7e-29;
Matches 122; Conservative 45; Mismatches 170; Indels 77; Gaps 10;

Qy 47 LPGYKYLGPFGNLDKGPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEPQERLQE 106
Db 134 LPGTNYVVGPNELQAGPPQSAVDSAARIHDFRYSQAKLGINPYTHWTVADELLKNIK 193
Qy 107 DTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKRPVEQSPQBPSSSGIGKTQQP 166
Db 194 ETGFQAQVVKDYF-----TLKGAAPVAHFQ-----GSLPEVP 226
Qy 167 AKRLNFGQTGDSSEVPDPQPLGEPPATPAAGVPTTMSGGGAPMADNNEGADGVGNASG 226
Db 227 AYNA-----SEKYPSTSVNSAEASTGA-----GGG-----GSNSVKS 259
Qy 227 NWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISSASTGASND-----NHYFGY 276
Db 260 MWSEGATFSANSVTCTFSRQFLIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGY 319
Qy 277 STPWGYFDENRPHCHFSRQDLINNNWGFPRKRLNFKLFNIQVKEVT--TNDGVTTIA 334
Db 320 STPWRYLDFNALNLFSPLEFQHLIENYGSIAPDALTVTIASEIAVKDVTDKTGGV-QVT 378
Qy 335 NNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTLNN-GSQAVG---- 389
Db 378 DSTTGRCLMLVDHEYKYPYVLGGQDTLAPELPIWVYFPPQYAYLTVGDVNTQGISGDSK 438
Qy 390 -----RSSFYCLEYFPQMLRTGNFTFSYTFEEVPPHSSYAHSQSLDRLMNPL 438
Db 438 KLAESAFYVLEHSSFQLLGTGTASMSYKFPVPVPPENLEGCSQH FYEMYNPL 492

RESULT 9
US-08-856-841-19
; Sequence 19, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 486
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 19:
US-08-856-841-19

Query Match 10.8%; Score 429.5; DB 3; Length 486;
Best Local Similarity 28.8%; Pred. No. 4.3e-29;
Matches 120; Conservative 45; Mismatches 169; Indels 83; Gaps 7;
QY 47 LPGAQVVKDYFTLKGAAAPVAHFQGLPEVPAYNASEKYPMSVTSVNSAEASTGAGGG 244
DB 125 LPGAQVVKDYFTLKGAAAPVAHFQGLPEVPAYNASEKYPMSVTSVNSAEASTGAGGG 244
QY 107 DTSGGNNLGRAVFOAK---KRVLEPLGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTG 163

Db 185 ETGFAQVVKDYFTLKGAAAPVAHFQGLPEVPAYNASEKYPMSVTSVNSAEASTGAGGG 244
QY 164 QPAKAKRLNFGTGDSVPDPQPLGEPATPAAVGPTTMAASGGAPWADNNEGADGVGN 223
Db 245 SNPVKSM----- 251
QY 224 ASGNWHCDSTWLGDRVITTTSTRTWALPTNNHLYKQISSASTGASND-----NHY 273
Db 252 ---WSEGATFSANSVTCTFSRQFLIPYDPEHHYKVFSAPASSCHNASGKAEKVCITSPI 307
QY 274 FGYSTPWGYFDENRFCHFSRDRWQRLNNWGFRRPKRLNFKLFNIQVKEVT--TNDGVT 331
Db 308 MGYSTPWRYLDFNALNLFSPLEFQHLIENYGSTAPDALTTTISEIAVKDVTDKTGGGV- 366
QY 332 TIANNLTSTVQVPSDSEYQLPYVLGSAHQGCLPPPPADVFPMIPQYGYLTLLNN-GSQAVG- 389
Db 367 QVTDSATGRCLMLVDHEYKYYPYVLGQGGDTLAPELPIWVYPPQYAYLTVGDVNTQGSG 426
QY 390 -----RSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPPFHSSYAHSSQSLDRMLNPL 438
Db 427 DSKCLASESAFYVLEHSSFQLLGTGTATMSYKFPVPENLEGCSQHPYEMYNPL 483
RESULT 10
US-08-856-841-20
Sequence 20, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:

US-08-856-841-20

Query Match 8.2%; Score 326; DB 3; Length 415;
Best Local Similarity 26.0%; Pred. No. 4e-20;
Matches 101; Conservative 66; Mismatches 180; Indels 42; Gaps 13;

QY 330 VTTIANNLSTVQVFSRQYLPVVLGSAHQCLPPFPADVFMIPQYGLTLNN-GSQAV 388
Db 1 MTMITPSLHACMLV--DHEVKYPPVVLGGQDITLAPBLPIWVFPFPQYAYLTGVDVNTQGI 58
QY 389 G-----RSSFYCLEYPPSQMLRTGNFTFSYTFEEVFFHSSYAHQSQSLDRLMNP LI 439
Db 59 SGDSKKLASEESAFVLEHSSFQLLGTGTASMSYKFPVPFPENLEGCSQHFYEMYNPL- 117
QY 440 DQYLYLNRNTQNSGSAQNKDILLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNS 499
Db 118 --YGSRLGVDPDTLGGDPKFRSL-----THEDHAIQPNFMFGPLVNSVSTKEGSSNTGA 170
QY 500 NFWTGASKYNLNGRESIINPG-TAMASHKDDKEDKFFPMGSMVIFGKESAGASNTALDNV 558
Db 171 GKALTGLSTGTSQNTIRSLRPGVSPQYHHWDTDKYVTGAINAISHGQTTYG----NAEDKE 227
QY 559 -----MITDEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWVQDRD 611
Db 228 YQGVGRFPNEKEQLKQLOGLNHMTY-----FPNKGTOQYTDQIE-RPLMVGSVWNNRA 280
QY 612 VYLQGIWAKIPHTDGHFHS-PLMGGFGLKNPPQILIKNTVPVPANPPAEFSATKFSF 670
Db 281 LHVESQLWSKIPNLDDSFKTQFAALGGWGLHQPPQIFLK--ILPESGPIGGIKSMGITT 338
QY 671 ITQYSTGQVSVIEWEL-QKENSQRWNPE 698
Db 339 LVQYAVGIMTMTFTKLGPRKATGRWNPO 367

RESULT 11

US-08-856-841-13
; Sequence 13, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING

JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:

Query Match 8.0%; Score 318; DB 3; Length 395;
Best Local Similarity 25.4%; Pred. No. 1.9e-19;
Matches 99; Conservative 61; Mismatches 167; Indels 62; Gaps 13;

Qy 330 VTTIANMLTSTVQVFSQSEYQLPVYLGSAHQGLPPFPADVPFMIPOYGYLTNN-GSQAV 388
Db 1 MTMITPSLHACMLV--DHEYKYPYVLGGQDTLAPLPWYPPQYAYLTVGDVNTQGI 58
Qy 389 G-----RSSFYCLEYFPSPMLRTGNFTFSYTFEEVPHSSYAHSQSLDRLMNP LI 439
Db 59 SGDSKKLASEESAFVLEHSSFQLLTGTGTASMSYKFPVPVPENLEGCSQH FYEMYNPL- 117
Qy 440 DQYLYLNRQNSQSAQNKDLESRGSPAGMSVQPKNLPGPCYRQQRVSKTKTDNNNS 499
Db 118 --YGRLGVPDPTLGGDPKFRSL-----THEDHAIQPNFMPGLVNSVSTKEGDSNTGA 170
Qy 500 NFTWTGASKYNLNGRESIINPG-TAMASHKDDKFFPMGVMIFGKESAGASNTALDNV 558
Db 171 GKALTGLSTGTSQNTSRISLRPGVPVSPYHHWTDKYVTGINAISHGQTTYG---NAEDKE 227
Qy 559 -----MTDEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRD 611
Db 228 YQGVGRFPNEKEQLKQLGLNMHTY-----FPNKGTOQYTDQIE-RPLMVGSVMNRRRA 280
Qy 612 VYLGQPIWAKIPHTDGHFHS-PLMGGFGLKNPPPPQILIKNTPVPANPPAEFSATKFASE 670
Db 281 LHYESQLWSKIPNLDDSFKTQFAALGGWGLHQPPPI-----F 318
Qy 671 ITQYSTGQVSVEIEWEL-QKENSKEWNPE 698
Db 319 LKQYAVGIMTVMTFKLGP RKATGRWNPQ 347

RESULT 12
US-08-856-841-14
; Sequence 14, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK

STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 264
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 14:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:

Qy 568 ATNPVATERFGTVAVNFQSSSTDBATGDVHAMGALPGMVWQDRDVYLQGPWAKIPHTDG 627
Db 233 QLOGLNMHTY-----FPNKGTOQYTDQIE-RPLMVGSVNRRALHYESQLWSKIPNLDD 285
Qy 628 HFHPS-PLMGCGFLKNPPPPQILIKNTFPVPANPPAEFSATKFAFITQYSTGQVSVEIEWE 686
Db 286 SFKTQPAALGGWGLHQPPPI-----FKYYHKVGQLEVLNQWE 323
Qy 687 LQKENS-----RWNP 697
Db 324 LLPFNPWELQLHLNWGP 341

RESULT 14
US-08-856-841-17
; Sequence 17, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 387
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A

; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; AUTHORS: WIDDOWS, D.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET
; VOLUME: I
; ISSUE:
; PAGES: 72 - 73
; DATE: 1975
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 17:
; PUBLICATION INFORMATION:
; AUTHORS: MANIATIS, T.
; AUTHORS: FRITSCH, E.F.
; AUTHORS: SAMBROOK, J.
; TITLE: MOLECULAR CLONING
; JOURNAL: COLD SPRING HARBOR, NY
; VOLUME:
; ISSUE:
; PAGES:
; DATE: 1982
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 15:
; PUBLICATION INFORMATION:
; AUTHORS: SMITH, D.B.
; AUTHORS: JOHNSON, K.S.
; TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
; TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
; TITLE: GLUTATHIONE S. TRANSFERASE
; JOURNAL: GENE
; VOLUME:
; ISSUE: 67
; PAGES: 31 - 40
; DATE: 1988
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 17:
; US-08-856-841-17

Query Match 7.1%; Score 284; DB 3; Length 387;
Best Local Similarity 28.0%; Pred. No. 1.7e-16;
Matches 82; Conservative 28; Mismatches 119; Indels 64; Gaps 6;
Qy 47 LPGYKYLGPFGNGLDKGEPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQE 106
Db 133 LPGTNYVGPNGNELQAGPPQSAVDSAARIHDFRYSQLAKLGINPYTHWTVADEELLKNIK 192
Qy 107 DTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPGKCRPVEQSPQEPDSSSGIGKTQQP 166
Db 193 ETGFQAQVVKDYF-----TLKGAAPVAHFQ-----GSLPEVP 225
Qy 167 AKRLNFGQTGDSSEVPDPQPLGEPPTPAAVCPPTTMASGGGAPMADNEGADGVGNASG 226
Db 226 AYNA-----SEKYPSTSVNSAEASTGA-----GGG-----GSNSVKS 258
Qy 227 NWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISSASTGASND-----NHYFGY 276
Db 259 MWSEGATFSANSVTCTFSRQFLIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGY 318
Qy 277 STPWGYFDNFRHFCHFSRQWLINNNWGFPPKRLNFKLFNIQVKEVTTNDG 329
Db 319 STPWRYLDFNALNLFFSPLEFQHLIENYGSYAPDALVTVTISEIAVKDVTDKTG 371

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: January 21, 2004, 15:59:46 ; Search time 36.5361 Seconds
(without alignments)
2602.280 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTERPRPIGTRYLTRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1107863 seqs, 158726573 residues 1107863
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					§	
Result No.	Score	Query Match	Length	ID	Description	
1	3251	100.0	599	21	AAV71168	Adeno-associated v
2	3251	100.0	736	21	AAV71167	Adeno-associated v
3	3229	99.3	736	22	AAB59847	AAV6 capsid protei
4	2906	89.4	534	21	AAV71169	Adeno-associated v
5	2832	87.1	736	22	AAB59846	AAV3B capsid prote
6	2815	86.6	736	22	AAB59845	AAV3A capsid prote
7	2764.5	85.0	735	22	AAG65792	Adeno-associated v
8	2764.5	85.0	735	22	AAM51508	Adeno-associated v
9	2764.5	85.0	735	22	AAB59844	AAV2 capsid protei

10	2764.5	85.0	735	23	AAU98974	Adeno-associated v
11	2759.5	84.9	598	22	AAG65793	Adeno-associated v
12	2759.5	84.9	598	22	AAM51509	Adeno-associated v
13	2759.5	84.9	598	23	AAU98975	Adeno-associated v
14	2738	84.2	734	22	AAB50326	Adeno-associated v
15	2481.5	76.3	533	22	AAG65794	Adeno-associated v
16	2481.5	76.3	533	23	AAU98976	Adeno-associated v
17	2477.5	76.2	533	22	AAM51510	AAV4 VP2 coat prot
18	1830.5	56.3	598	19	AAW46313	Adeno-associated v
19	1830.5	56.3	598	24	ABG73940	AAV4 VP1 capsid pr
20	1830.5	56.3	734	19	AAW46308	Adeno-associated v
21	1830.5	56.3	734	24	ABG73935	Duck parvovirus ca
22	1719	52.9	736	18	AAW19000	Barbary duck parvo
23	1701	52.3	732	16	AAR85385	Adeno associated v
24	1700.5	52.3	588	21	AAV58161	Adeno-associated v
25	1700.5	52.3	588	23	AAU11406	Adeno-associated v
26	1700.5	52.3	724	21	AAV58160	Adeno associated v
27	1700.5	52.3	724	23	AAU11405	Adeno-associated v
28	1690.5	52.0	544	19	AAW46314	AAV4 VP3 coat prot
29	1690.5	52.0	544	24	ABG73941	Adeno-associated v
30	1676	51.6	534	16	AAR85386	Barbary duck parvo
31	1665	51.2	532	21	AAV58162	Adeno associated v
32	1665	51.2	532	23	AAU11407	Adeno-associated v
33	772	23.7	756	21	AAV71231	Capsid protein enc
34	497.5	15.3	781	16	AAW08986	Human parvovirus V
35	497.5	15.3	781	20	AAV23227	Erythrovirus V9 VP
36	493	15.2	781	24	ABP57263	Human parvovirus B
37	490	15.1	554	16	AAW08987	Human parvovirus V
38	490	15.1	554	20	AAV23230	Erythrovirus V9 VP
39	486	14.9	554	24	ABP57264	Human parvovirus B
40	486	14.9	554	24	ABP57267	Human parvovirus B
41	484	14.9	781	24	ABP57266	Human parvovirus B
42	476.5	14.7	543	12	AAR13405	Parvo virus B19 VP
43	310	9.5	264	12	AAR13407	Parvo virus B19 PA
44	295.5	9.1	370	12	AAR13406	Parvo virus B19 PA
45	283	8.7	202	23	AAU83481	Adeno-associated v

ALIGNMENTS

RESULT 1
AAV71168
ID AAV71168 standard; Protein; 599 AA.
XX
AC AAV71168;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP2.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP2.
XX
OS Adeno associated virus serotype 1.
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00778.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for

PT preparation of medicament for delivery of a transgene to a host -
XX
PS Claim 7; Page 93-95; 108pp; English.

The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is an AAV-1 cap protein VP2 which is useful in the production of recombinant viral vector for gene delivery.

Sequence 599 AA;

Query Match	100.0%	Score 3251;	DB 21;	Length 599;
Best Local Similarity	100.0%;	Pred. No. 1.5e-254;		
Matches 599;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSVSDPQPLGEPPATPAA 60

Dd 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSVSDPQPLGEPPATPAA 60

Qy	61	VGPTT	MAS	GGG	GAP	PMAD	NN	EG	AD	GV	GN	AS	GN	W	H	CD	ST	W	L	G	D	R	V	I	T	T	S	T	R	T	W	A	L	P	T	Y	N	N	H	L	Y	120
Qy	61	VGPTT	MAS	GGG	GAP	PMAD	NN	EG	AD	GV	GN	AS	GN	W	H	CD <td>ST</td> <td>W</td> <td>L</td> <td>G</td> <td>D</td> <td>R</td> <td>V</td> <td>I</td> <td>T</td> <td>T</td> <td>S</td> <td>T</td> <td>R</td> <td>T</td> <td>W</td> <td>A</td> <td>L</td> <td>P</td> <td>T</td> <td>Y</td> <td>N</td> <td>N</td> <td>H</td> <td>L</td> <td>Y</td> <td>120</td>	ST	W	L	G	D	R	V	I	T	T	S	T	R	T	W	A	L	P	T	Y	N	N	H	L	Y	120
db	61	VGPTT	MAS	GGG	GAP	PMAD	NN	EG	AD	GV	GN	AS	GN	W	H	CD <td>ST</td> <td>W</td> <td>L</td> <td>G</td> <td>D</td> <td>R</td> <td>V</td> <td>I</td> <td>T</td> <td>T</td> <td>S</td> <td>T</td> <td>R</td> <td>T</td> <td>W</td> <td>A</td> <td>L</td> <td>P</td> <td>T</td> <td>Y</td> <td>N</td> <td>N</td> <td>H</td> <td>L</td> <td>Y</td> <td>120</td>	ST	W	L	G	D	R	V	I	T	T	S	T	R	T	W	A	L	P	T	Y	N	N	H	L	Y	120

Qy	121	KQISSASTGASNDNHYFGYSTPWGYFDENRPFCHFS	PRDWQRLINNNWGFRKLNFKLF	180
db	121	KQISSASTGASNDNHYFGYSTPWGYFDENRPFCHFS	PRDWQRLINNNWGFRKLNFKLF	180

Qy	181	NIQKEVTNDGVTTIANNLTSITVQVFSDSEYQLPYVLGSAHQGCLPPFPADVEMIPQYG	240
Qy	181	NIQKEVTNDGVTTIANNLTSITVQVFSDSEYQLPYVLGSAHQGCLPPFPADVEMIPQYG <td>240</td>	240
pb	181	NIQKEVTNDGVTTIANNLTSITVQVFSDSEYQLPYVLGSAHQGCLPPFPADVEMIPQYG <td>240</td>	240

Qy

241 YLTUNNGSQAVGRSSFYCYLEYFPQSMLRTGNFTFSYTFEEVPFHSSYAHQSLSLDRLMNP 300
|||||

Dh

241 YLTUNNGSQAVGRSSFYCYLEYFPQSMLRTGNFTFSYTFEEVPFHSSYAHQSLSLDRLMNP 300
|||||

	LIDQLYYLNR	TQNSGSAQNKDL	FSRGSAGMSVQPKWLPGPCYRQRVSKTKTDNN	360
Qy				
	LIDQVIYYI	NPTONSGSAQNKI	LFSPGSDAGMSVQPKWLPGPCYRQRVSKTKTDNN	360
Dh				

Qy	361	NSNFTWTGASKYNLNGRESII	INPGTAMASHKODEDKFFPM	SGVMI	FGKESAGASNTALDN	420
Dh	361	NSNFTWTGASKYNI	NGRESI	INPGTAMASHKONDEKFFPM	SGVMI	FGKESAGASNTALDN
						420

[illegible][illegible]

QY 541 QVSVEIENELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599

RESULT 2
AAAY711167
ID AAAY711167 standard: Protein: 736 AA:

AA AAY71167;
AC

DT 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 capsid protein VP1.

Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
cap protein; recombinant viral vector; gene delivery; gene therapy;
vaccine; transgene; VP1.

Adeno associated virus serotype 1.

PN WO200028061-A2.

PD 18-MAY-2000.

PF 02-NOV-1999; 99WO-US25694.

PR 05-NOV-1998; 98US-0107114.

PA (UYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Xiao W;

DR WPI; 2000-376571/32.

XX

PT preparation of medicament for delivery of a transgene to a host -

PS Claim 7; Page 87-90; 108pp; English.

The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transducing host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is an AAV-1 cap protein VP1 which is useful in the production of recombinant viral vector for gene delivery.

AA	Sequence	736 AA;
SO		

Query Match 100.0%; Score 3251; DB 21; Length 736;
Best Local Similarity 100.0%; Pred. No. 2e-254;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TAPGKCRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVPDPQPLGEPPATPAA 60
|||
QY 138 TAPGKCRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSVPDPQPLGEPPATPAA 197
|||

Qy	61	VGPTT	ASGGG	APMAD	NEGAD	GVGN	ASGN	WHCD	STWL	GD	VIT	TTST	RT	WT	AL	PT	YN	NH	LY	120
ph	198	VGPTT	ASGGG	APMAD	NEGAD	GVGN	ASGN	WHCD	STWL	GD	VIT	TTST	RT	WT	AL	PT	YN	NH	LY	257

Qy	121	KQISSASTGASNDNHYFGYSTPWGYFD	FNRFHCHFS	PRDWORLINNNWGF	PKRLNFKLF	180
nb	258	KQISSASTGASNDNHYFGYSTPWGYFD	FNRFHCHFS	PRDWORLINNNWGF	PKRLNFKLF	317

Qy	181	NIOVKVTTNDGVTTIANNLTSTVOVFSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYG	240
yb	219	NIOVKVTTNDGVTTIANNLTSTVOVFSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYG	377

QY 241 YLT L N G S Q A V G R S S F Y C L E Y P P S Q M L R T G N N F T F S Y T F E E V P F H S S Y A H S Q S L D R L M N P 300

QY 301 LIDQYLYLNRNTONQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360

QY	361	NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGVMIFGKESAGASNTALDN	420
Db	498	NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGVMIFGKESAGASNTALDN	557
QY	421	VMITDEEEIKATNPVATERFGTVAVNPFQSSSTDPA TGDVHAMGALPGMWQDRDVLQGP	480
Db	558	VMITDEEEIKATNPVATERFGTVAVNPFQSSSTDPA TGDVHAMGALPGMWQDRDVLQGP	617
QY	481	IWAKIPHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTG	540
Db	618	IWAKIPHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTG	677
QY	541	QVSVEIEWELQKENSKRWNPEVOYTSNYAKSANVDFTV DNNGLYTEPRPIGTRYLTRPL	599
Db	678	QVSVEIEWELQKENSKRWNPEVOYTSNYAKSANVDFTV DNNGLYTEPRPIGTRYLTRPL	736

Qy	61	VGPTT	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLY	120
Db	198	VGPTT	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLY	257
Qy	121	KQISS	ASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNWNWGFPRKRLNFKLF	180
Db	258	KQISS	ASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNWNWGFPRKRLNFKLF	317
Qy	181	NIQVKE	VTTNDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPOYG	240
Db	318	NIQVKE	VTTNDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPOYG	377
Qy	241	YLTNNG	SQAVGRSSFYCLEYPPSQMLRTGNFTFSYTFEEVPHSSYAHQSLSLDRLMNP	300
Db	378	YLTNNG	SQAVGRSSFYCLEYPPSQMLRTGNFTFSYTFEEVPHSSYAHQSLSLDRLMNP	437
Qy	301	LIDQYL	YLNLRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGCYRQQRVSKTKTDNN	360
Db	438	LIDQYL	YLNLRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGCYRQQRVSKTKTDNN	497
Qy	361	NSNFTW	TGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN	420
Db	498	NSNFTW	TGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN	557
Qy	421	VMITDEE	EIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRD VYLQGP	480
Db	558	VMITDEE	EIKATNPVATERFGTVAVNLQSSSTD PATGDVHVMGALPGMWQDRD VYLQGP	617
Qy	481	IWAKIPH	TDGHFHPSLMCGGFLKNPPQIILIKNTVPANPPAEFSATKFASFITQYSTG	540
Db	618	IWAKIPH	TDGHFHPSLMCGGFLKHPPQIILIKNTVPANPPAEFSATKFASFITQYSTG	677
Qy	541	QVSVEI	EWELQENSKRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL	599
Db	678	QVSVEI	EWELQENSKRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL	736

XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 cap protein vp3 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 534 AA;

Query Match 89.4%; Score 2906; DB 21; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-226;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 MASGGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 125
Db 1 MASGGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 60

Qy 126 ASTGASNDNHYFGYSTPWGYFDNFRFHCFSPRDWQRLINNNGWFRPKRLNFKLFNQVK 185
Db 61 ASTGASNDNHYFGYSTPWGYFDNFRFHCFSPRDWQRLINNNGWFRPKRLNFKLFNQVK 120

Qy 186 EVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 245
Db 121 EVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 180

Qy 246 NGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 305
Db 181 NGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 240

Qy 306 LYTLNRTQNSGSAQNKOLLFSRGSAGMSVQPKNLPGPCYRQQRVSKTKTDNNNSNFT 365
Db 241 LYTLNRTQNSGSAQNKOLLFSRGSAGMSVQPKNLPGPCYRQQRVSKTKTDNNNSNFT 300

Qy 366 WTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMSGVMI FGKESAGASNTALDNVMITD 425
Db 301 WTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMSGVMI FGKESAGASNTALDNVMITD 360

Qy 426 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGVMWQDRDVLQGPWAKI 485
Db 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGVMWQDRDVLQGPWAKI 420

Qy 486 PHTDGHFHPSPLMGGFGLKNPPPOILLIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 545
Db 421 PHTDGHFHPSPLMGGFGLKNPPPOILLIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480

Qy 546 IEWELQKENS KRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 481 IEWELQKENS KRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534

RESULT 5
AAB59846
ID AAB59846 standard; Protein; 736 AA.
XX
AC AAB59846;
XX
DT 28-MAR-2001 (first entry)
XX
DE AAV3B capsid protein VP1.
XX
KW AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.
XX

OS Adeno associated virus.
XX US6156303-A.
PN
XX
PD 05-DEC-2000.
XX
PF 11-JUN-1997; 97US-0873168.
XX
PR 11-JUN-1997; 97US-0873168.
XX (UNIW) UNIV WASHINGTON.
PA
XX Russell DW, Rutledge EA;
PI
XX WPI; 2001-060164/07.
DR
XX
PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein vp1 of one such serotype (AAV3B).
CC AAV3B can be used to construct AAV viral vectors for use in gene therapy
CC for a range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX
SQ Sequence 736 AA;

Query Match 87.1%; Score 2832; DB 22; Length 736;
Best Local Similarity 85.7%; Pred. No. 1.8e-220;
Matches 514; Conservative 34; Mismatches 50; Indels 2; Gaps 2;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTQQQPAKQKLNFGQTGDSSEVPDPQPLCEPPATPAA 60
Db 138 TAPGKKRPVDQSPQEPDSSSGVGKSGKQPKARKKLNFGQTGDSSEVPDPQPLCEPPAPTS 197

Qy 61 VGPTTMASGGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLY 120
Db 198 LGSNTMASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLY 257

Qy 121 KQISSASTGASNDNHYFGYSTPWGYFDNFRFHCFSPRDWQRLINNNGWFRPKRLNFKLF 180
Db 258 KQISSQS -GASNDNHYFGYSTPWGYFDNFRFHCFSPRDWQRLINNNGWFRPKKLSFKLF 316

Qy 181 NIQVKEVTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQY 240
Db 317 NIQVKEVTQNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQY 376

Qy 241 YLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNP 300
Db 377 YLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNP 436

Qy 301 LIDQYLYLNRTO -NQS GSAQNKDLLFSRGSAGMSVQPKNLPGPCYRQQRVSKTKTDN 359
Db 437 LIDQYLYLNRTOGTTSGTTNQSRLLFSSQAGPQSNLSQARNWLPGPCYRQQRLSKTANDN 496

Qy 360 NNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMSGVMI FGKESAGASNTALD 419
Db 497 NNSNFPWTAASKYHLNGRDSLVPNGPAMASHKDDDEKFFPMHGNLIFGKEGTTASNABLD 556

Qy 420 NWMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGVMWQDRDVLQ 479
Db 557 NWMITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTRTVNDQGALPGVMWQDRDVLQ 616

Qy 480 PIWAKIPHTDGHFHPSPLMGGFGLKNPPPOILLIKNTVPANPPAEFSATKFAFITQYST 539
Db 617 PIWAKIPHTDGHFHPSPLMGGFGLKHPPPOIMIKNTVPANPPPTTFSPAKPASFITQYST 676

a cell with RVV having (I) such that (I) binds to an attachment site on the cell surface and permits the vector to enter the cell efficiently. A pharmaceutical composition comprising RVV with (I) containing a transgene sequence associated with a disease or a disorder such that expression of the transgene would result in amelioration of the disease or disorder such as inherited neurological and metabolic diseases e.g. lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polyneuropathy, Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic fibrosis, diabetes, diseases associated with hormone deficiencies, retinoblastoma and various types of neoplastic cells which include tumours especially central nervous system tumours, neoplasms, carcinomas, sarcomas, leukemias and lymphoma. The present sequence represents the adeno-associated virus 2 (AAV-2) major coat protein VP1. AAV-2 sequences are used in the construction of a chimeric vector.

Adeno-associated virus; AAV; VP1; capsid; virus-like particle;
nuclear localisation signal; VP3.
Adeno associated virus.
JP2001169777-A.
26-JUN-2001.
30-JUL-1999; 99JP-0249140.
30-JUL-1999; 99JP-0249140.
(HAND//) HANDA H.
WPI; 2001-599854/68.
New virus-like particles from VP3 capsid protein of adeno-associated virus, comprise a peptide containing a nucleus-shifting signal connected to its N-terminal -
Disclosure; Page 10-13; 33pp; Japanese.
The present sequence is provided in a specification relating to a virus-like particle-forming protein, and to a peptide containing a nuclear-shifting signal at its N-terminus that can form virus-like particles by shifting to the nucleus of the animal cell in which it is expressed. The method is used for forming virus-like particles from the VP3 capsid protein of adeno-associated virus (AAV)

Db 677 QVSVEIEWELQKENS KRWNP EIQYTSNYKSVNVDFTVD TNGVYSEPRPIGTRYLTRNL 735

RESULT 9
AAB59844
ID AAB59844 standard; Protein; 735 AA.
XX AAB59844;
AC AAB59844;
XX 28-MAR-2001 (first entry)
DT AAV2 capsid protein VP1.
DE AAV2;
XX AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.
XX Adeno associated virus.
OS US6156303-A.
XX 05-DEC-2000.
PD 11-JUN-1997; 97US-0873168.
XX 11-JUN-1997; 97US-0873168.
PR (UNIW) UNIV WASHINGTON.
XX Russell DW, Rutledge EA;
PI WPI; 2001-060164/07.
XX Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -
XX Claim 7; Fig 2; 50pp; English.
XX The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV2). AAV2
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX Sequence 735 AA;
SQ

Query Match 85.0%; Score 2764.5; DB 22; Length 735;
Best Local Similarity 82.8%; Pred. No. 5.1e-215;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDSBSPDPQPLGEPATPAA 60
Db 138 TAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQPPA 197

Qy 61 VGPTTMASCGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLY 120
Db 198 LGTNTMATSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNHLY 257

Qy 121 KOISSASTGASNDNHFGYSTPWGYFDNRFCHFPSRDWQRLNNNWGFRPKRLNFKLF 180
Db 258 KOISSQS-GASNDNHFGYSTPWGYFDNRFCHFPSRDWQRLNNNWGFRPKRLNFKLF 316

Qy 181 NIQVKEVTTNDGVTTIANNLTSTQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPQYG 240
Db 317 NIQVKEVTQNDGTTIANNLTSTQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYG 376

Qy 241 YLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNTFSYTFEEVFFHSSYAHQSQSLDRLMNP 300

Db 377 YLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNTFSYTFEDVFFHSSYAHQSQSLDRLMNP 436

Qy 301 LIDQYLYLNRNTQNQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQORVSKTKTDNN 360
Db 437 LIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDOSRNWLPGPCYRQORVSKTSADNN 496

Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMGSMVIFGKESAGASNTALDN 420
Db 497 NSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKDEKFFQSGVLIIFGKQGSEKTNVDIEK 556

Qy 421 VMITDEEEIKATNPVATERFCTVAVNFQSSSTDPA TGDVHAMGALPGMVWQDRDVYLOGP 480
Db 557 VMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLOGP 616

Qy 481 IWAKIPHTDGHFSPPLMGFGGLKNPPQILIKNTVPANPPAEFSATKFAFITOYSTG 540
Db 617 IWAKIPHTDGHFSPPLMGFGGLKHPPPQILIKNTVPANPSTTFSAAKFASFITOYSTG 676

Qy 541 QVSVEIEWELQKENS KRWNP EVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 677 QVSVEIEWELQKENS KRWNP EIQYTSNYKSVNVDFTVD TNGVYSEPRPIGTRYLTRNL 735

RESULT 10
AAU98974
ID AAU98974 standard; Protein; 735 AA.
XX AAU98974;
AC AAU98974;
XX 05-NOV-2002 (first entry)
DT Adeno-associated virus 2 (AAV2) vector, VP1 capsid protein.
XX Adeno-associated virus 2 (AAV2) vector, VP1 capsid protein.
DE Adeno-associated virus 2 vector; AAV2; cancer; VP1 capsid;
KW heparin-sulphate proteoglycan; vaccine; immune response;
KW ovarian cancer.
XX Adeno-associated virus 2.
OS WO200253703-A2.
XX 11-JUL-2002.
XX 04-JAN-2002; 2002WO-US00152.
PF 05-JAN-2001; 2001US-260124P.
PR (CHIL-) CHILDRENS HOSPITAL INC.
PA Bartlett JS;
XX WPI; 2002-583608/62.
XX N-PSDB; ABK89694.
DR New adeno-associated virus vector comprises a biotinylated capsid or
PT capsid protein with an amino acid insertion in the VP1 capsid, useful
PT as a vaccine or for transferring a therapeutic peptide to a cancer cell
PT -
XX Claim 1; Page 49-51; 57pp; English.
XX The invention relates to an adeno-associated virus (AAV) vector (I)
CC comprising a biotinylated capsid or capsid protein (II) with an amino
CC acid insertion following the capsid amino acid at position 139, 161, 588
CC or 657 in the VP1 capsid. The AAV vector comprises a capsid protein
CC containing one or more amino acid insertions that ablate the ability of
CC the vector to bind heparin-sulphate proteoglycan and allow the vector to
CC use a cellular receptor not used by wild type AAV. Modified (I) are
CC useful as vaccines to elicit immune responses to amino acids, where the
CC response can be protective and/or therapeutic. (I) may be used to
CC transfer a therapeutic peptide to a cancer cell, particularly to an
CC ovarian cancer cell. The present sequence represents the adeno-associated

Db 616 IWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTG 675
Qy 541 QVSVEIEWELOKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 599
Db 676 QVSVEIEWELOKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 734

RESULT 15
AAG65794
ID AAG65794 standard; Protein; 533 AA.
XX
AC AAG65794;
XX
DT 11-FEB-2002 (first entry)
XX
DE Adeno-associated virus 2 (AAV-2) major coat protein VP3.
XX
DE Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytostatic;
KW inverted terminal repeat; nootropic; neuroprotective; antianemic; ITR;
KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV;
KW major coat protein; AAV-2; VP3.
XX

OS Adeno-associated virus 2.
XX
XX WO200168888-A2.
PN
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07927.
XX
PR 14-MAR-2000; 2000US-189110P.
XX
PA (NEUR-) NEUROLOGIX INC.
XX
PI Xiao W, During MJ;
XX
DR WPI; 2001-596912/67.
DR N-PSDB; AAI66974.
XX
PT Recombinant viral vector useful in improving gene therapy in a subject,
PT and for increasing efficiency of entry into a cell, comprises a
PT chimeric capsid having one non-native amino acid sequence and a desired
PT transgene -

XX
PS Disclosure; Page 51; 53pp; English.
XX
CC The invention provides a recombinant viral vector (RVV) comprising a
CC chimeric capsid (I) having at least one non-native amino acid sequence,
CC derived from a capsid protein domain of parvovirus (II), a virus (III),
CC or their combination, and a transgene flanked 5' and 3' by inverted
CC terminal repeat (ITR) sequences, derived from (II), (III), or their
CC combination. The RVV is useful for improving gene therapy in a subject
CC with a disorder, and for increasing the efficiency of entry into a cell,
CC which involves producing (I) encapsulating a viral vector, and contacting
CC a cell with RVV having (I) such that (I) binds to an attachment site on
CC the cell surface and permits the vector to enter the cell efficiently.
CC A pharmaceutical composition comprising RVV with (I) containing a
CC transgene sequence associated with a disease or a disorder such that
CC expression of the transgene would result in amelioration of the disease
CC or disorder such as inherited neurological and metabolic diseases e.g.
CC lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polynuropathy,
CC Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood
CC e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic
CC fibrosis, diabetes, diseases associated with hormone deficiencies,
CC retinoblastoma and various types of neoplastic cells which include
CC tumours especially central nervous system tumours, neoplasms, carcinomas,
CC sarcomas, leukemias and lymphoma. The present sequence represents the
CC adeno-associated virus 2 (AAV-2) major coat protein VP3. AAV-2 sequences
CC are used in the construction of a chimeric vector.

XX Sequence 533 AA;

SQ

Query Match 76.3%; Score 2481.5; DB 22; Length 533;
Best Local Similarity 83.3%; Pred. No. 2.8e-192;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;
Qy 66 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 125
Db 1 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDGVITTTSTRTWALPTYNNHLYKQISS 60
Qy 126 ASTGASNDNHYFGYSTPWGYFDENRFHCHFSRDRWQRLINNNWGFPRKRLNFKLFNIQVK 185
Db 61 QS-GASNDNHYFGYSTPWGYFDENRFHCHFSRDRWQRLINNNWGFPRKRLNFKLFNIQVK 119
Qy 186 EVTTNDGVTTIANNLTSTVQVFSDESQYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTILN 245
Db 120 EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTILN 179
Qy 246 NGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEEVFFHSSYAHSQSLDRLMPLIDQY 305
Db 180 NGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEDVPFFHSSYAHSQSLDRLMPLIDQY 239
Qy 306 LYLYNRTONQSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 365
Db 240 LYLSRTNTPSGTTTQSRLLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 299
Qy 366 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNVMITD 425
Db 300 WTGATKYHLNGRDSLVPNGPAMASHKODEEKFFPQSGVLIFGKQSEKTNVDIEKVMITD 359
Qy 426 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKI 485
Db 360 EEEIRTTNPVATEQYGSVSTNLQRGNRQAATAADVNTQGVLPGMVWQDRDVLQGPWAKI 419
Qy 486 PHTDGHFHPSPMLMGGFGLKNPPPPQILIKNTVPANPPAEPFSAKFASFITQYSTQVSV 545
Db 420 PHTDGHFHPSPMLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTQVSV 479
Qy 546 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 599
Db 480 IEWELQKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 533

Search completed: January 21, 2004, 16:04:19
Job time : 37.5361 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:00:42 ; Search time 14.1017 Seconds
(without alignments)
4084.983 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177.5	67.0	504	1 VCPV3A	coat protein - ade
2	1701	52.3	732	2 S52210	coat protein VP1 -
3	556	17.1	673	1 VCPVB5	coat protein VP1 -
4	495	15.2	781	1 VCPV19	coat protein VP1 -
5	258.5	8.0	723	1 VCPVPP	coat protein VP1 -
6	254.5	7.8	729	1 VCPVNA	coat protein VP1 -
7	254.5	7.8	729	1 A60006	coat protein VP1 -
8	244	7.5	722	1 VCPVV2	coat protein VP1 -
9	234	7.2	587	1 B44276	coat protein VP1 -
10	216	6.6	716	1 VCPV2M	coat protein VP1 -
11	216	6.6	727	1 VCPV1F	coat protein VP1 -
12	214.5	6.6	718	1 VCPVIM	coat protein VP1 -
13	214	6.6	722	1 VCPVME	coat protein VP1 -
14	209	6.4	727	1 VCPVFP	coat protein VP1 -
15	208	6.4	737	1 VCPVCD	coat protein VP1 -
16	204	6.3	748	1 VCPVCP	coat protein VP1 -
17	198	6.1	584	2 S49594	capsid protein VP2
18	192	5.9	722	1 VCPVCN	coat protein VP1 -
19	191	5.9	702	1 VCPVAP	coat protein VP1 -
20	143.5	4.4	648	2 S50856	whn protein - rat
21	129	4.0	1072	2 A86827	hypothetical prote
22	123.5	3.8	690	2 AB0124	probable TonB-depe
23	123	3.8	1142	2 T37455	enamelin precursor
24	122.5	3.8	880	1 SYBSVS	valine-tRNA ligase
25	121	3.7	635	2 F96660	protein F2K11.10 l
26	121	3.7	1113	2 S28925	nuclear pore compl
27	120.5	3.7	931	2 T49710	related to glucan
28	119	3.7	667	2 A41311	transcription fact
29	118.5	3.6	1180	2 E86719	hypothetical prote

30	118	3.6	739	2 T52289	probable transketo
31	116.5	3.6	642	1 S34416	transcription fact
32	116.5	3.6	1338	2 T30565	MAP kinase kinase
33	115.5	3.6	345	1 G97024	probable phosphoes
34	114	3.5	956	2 T08144	myrosinase-binding
35	114	3.5	2271	2 F90073	hypothetical prote
36	113.5	3.5	2817	2 B97033	uncharacterized pr
37	113	3.5	1071	2 E85343	hypothetical prote
38	112.5	3.5	666	2 I52648	class A helix-loop
39	112.5	3.5	3078	2 T28432	variant-specific s
40	111	3.4	1777	2 T34369	hypothetical prote
41	110.5	3.4	655	1 ALKBG	cyclomaltodextrin
42	110.5	3.4	1742	2 T17120	cellulase (EC 3.2.
43	110	3.4	868	2 A82515	conserved hypothet
44	110	3.4	1296	2 C82521	hemolysin-type cal
45	110	3.4	2825	2 T14271	Doc4 protein, stre

ALIGNMENTS

RESULT 1

VCPV3A

coat protein - adeno-associated virus type 2

C;Species: adeno-associated virus type 2

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Jul-1999

C;Accession: A03698

R;Srivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A;Reference number: A03694; MUID:83164299; PMID:6300419

A;Accession: A03698

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-504 <SRI>

A;Cross-references: EMBL:J01901; NID:g209616; PIDN:AAA42376.1; PID:g209621; EMBL:M12405;

C;Superfamily: adeno-associated virus coat protein

C;Keywords: coat protein

Query Match 67.0%; Score 2177.5; DB 1; Length 504;
Best Local Similarity 82.5%; Pred. No. 8.9e-142;
Matches 392; Conservative 33; Mismatches 47; Indels 3; Gaps 2;

Qy	66	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS	125
Db	1	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLYKQISS	60
Qy	126	ASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWORLNNWGFPRKRLNFKLFNIQVK	185
Db	61	QS-GASNDNHFGYSTPWGYFDNRFHCHFSPRDWORLNNWGFPRKRLNFKLFNIQVK	119
Qy	186	EVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTLN	245
Db	120	EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVFVMPQYGYLTIN	179
Qy	246	NGSQAVGRSSFYCYLPPSQMLRTGNNTFSYTPBEEVPPFHSSYAHSQSLDRLMNPILDOY	305
Db	180	NGSQAVGRSSFYCYLPPSQMLRTGNNTFSYTPFDVPPFHSSYAHSQSLDRLMNPILDOY	239
Qy	306	LYYLNRTQNQSGSAQNKDLLFSRGS PAGMSVQPKWMLPGPCYRQQRVSKTKTDNNNSNPT	365
Db	240	LYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS	299
Qy	366	WTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGVMIFGKESAGASNTALDNVMTD	425
Db	300	WTGATKYHLNGRDSLVP--AMASHKDDDEKFFPQSGVLIFGKQSEKTNVIEKVMITD	357
Qy	426	EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDQVYLOQGIWAKI	485
Db	358	EEEIGTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDQVYLOQGIWAKI	417
Qy	486	PHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPVNPAPAEFSATKFASFITQYSTG	540

Db 418 PHTDGHFHPSPMLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTG 472

RESULT 2
S52210
coat protein VP1 - muscovy duck parvovirus
N;Alternate names: VP1 protein
C;Species: muscovy duck parvovirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000
C;Accession: S52210
R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A;Reference number: S52209
A;Accession: S52210
A;Molecule type: DNA
A;Residues: 1-732 <ZAD>
A;Cross-references: EMBL:X75093; NID:g609091; PIDN:CAA52984.1; PID:g609093
A;Experimental source: strain FM
C;Genetics:
A;Gene: VP1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match 52.3%; Score 1701; DB 2; Length 732;
Best Local Similarity 53.0%; Pred. No. 7.8e-109;
Matches 325; Conservative 83; Mismatches 165; Indels 40; Gaps 11;

QY 6 KRPVEQSPQEPDSSGIGK-TGQQPAKRLNFGQTGDSVDPDPQLG-----E 53
Db 141 EEPVNTAPAKKSS---GKLTDHDPVKKPKLSE----ENSPSPNSGGEASAAATEGSE 192

QY 54 PPATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALP 113
Db 193 PVAAP-----NMAEGSGAMGDSAGGADGVGNASGNWHCDSQWLGDVTITKTRTWLP 246

QY 114 TYNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSRPDQWRQLNNNWGFRPK 173
Db 247 SYNHHYQAITSGTNPDSN-TQYAGYSTPWGYFDNRFCHFSRPDQWRQLNNHGWIRPK 305

QY 174 RLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADV 233
Db 306 ALKPKIFNVQKEVTTQDQTKTIANNLTSTIIFTDNEHQLPYVLGSATEGTMPPFPADV 365

QY 234 FMIPQYGYLTN---NGSQAVGRSSFYCLEYFPQMLRTGNFTFSYTFEEVPHSSYAH 290
Db 366 YALPQYGYCTMHTNQSGARFNDRSAFYCLEYFPQMLRTGNFTFSYTFEEVPHSMFAH 425

QY 291 SQSLDRLMNPLIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKWLPQPCYRQQ 350
Db 426 SQDLRLMNPLLDQYLVNFSEV-NGGRNAQ-----FKKAVK3AFGAMGRNWLPGPKLLDQ 479

QY 351 RVSKTK--TDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGVMIFGK 408
Db 480 RVRAYSGGTDNYANWSIWSKGNKVFLLKDRYLLQPGPVATHTEDQASSVPAQNIIGIAK 539

QY 409 E--SAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALP 466
Db 540 DPYRSGSTLAGISDIMVTDEQEIAPTNGVWRPYGLITVTNEQNTTTAPTNAELEVLGALP 599

QY 467 GMVWODRDVYLQGPWAKIPHDTDGHFHPSPMLMGGFGLKHPPPQILIKNTVPANPPAEFS 526
Db 600 GMVWQNRDIYLGQPIWAKIPKTDGKPHPSPNLGGFGLHNPFPQVFIKNTVPADPPLEYV 659

QY 527 ATKFASFITQYSTGOVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDEFTVDNNGLYTE 586
Db 660 NQKWSYITQYSTGQCTVEMVWELRKENSKRWNPEIQFTSNFGNRTSTMTFAPNETGGYVE 719

QY 587 PRPIGTRYLTRPL 599
Db 720 DRLIGTRYLTQNL 732

VCPVB5
coat protein VP1 - bovine parvovirus
N;Contains: coat protein VP2
C;Species: bovine parvovirus
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C;Accession: A26104
R;Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J. Virol. 60, 1085-1097, 1986
A;Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A;Reference number: A26104; MUID:87061184; PMID:3783814
A;Accession: A26104
A;Molecule type: DNA
A;Residues: 1-673 <CHE>
A;Cross-references: EMBL:M14363; NID:g333454; PIDN:AAB59847.1; PID:g808805
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 17.1%; Score 556; DB 1; Length 673;
Best Local Similarity 25.0%; Pred. No. 2.7e-30;
Matches 160; Conservative 98; Mismatches 239; Indels 144; Gaps 19;

QY 1 TAPGKKRPVEQSPQEPDSSGIGKTGQQPAKRLNFGQTGDSVDPDPQLGEP--PATP 58
Db 100 TSKGGDRALKRKLKLYFARSNKGAKKANREPAPSTSNQONMEVSNIDPNDEAGNQPIELATR 159

QY 59 AAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNH 118
Db 160 SVVSGSGVGGG-----RGGSGVGYSTGGTGTGFSENIIVVTKTRQFICIDKNGH 211

QY 119 LYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSRPDQWRQLNNNWGFRPKRLNFK 178
Db 212 LYKS-EVLNTGDTAHRQY-AITTPWSYFNFNQYSSHSFPNDWQHLVNDYERFRPKAMIVR 269

QY 179 LFNIOVKEVTTNDGVTTIANN-LTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIP 237
Db 270 VYNLQIKQIMTDGAMGTVYNNDLTAGMHIFCDGHRYPYVQHPWDDQCMPELPSIWELP 329

QY 238 QYGYLT-----NGSQAVGRS-----SFYCLEYFPQMLRTGNFTFSYTFEEVPHS 286
Db 330 QYAYIPAPISVVDNNTTNTVEEHLKGVFLYMLENSDHEVLNRG----- 373

QY 287 SYAHSQSLDRLMNPLIDQYLYLNRNTQN-----QSGSAQNKDLLFSRGSP----- 331
Db 374 -----RIYRIYIQLWRLMRDRKQHHIQHASDDVQSTGQKQKNLLIORTKQPNKORF 424

QY 332 AGMSVQPKWLPQPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHK 391
Db 425 QNAALRTSNWMSGP-----GIARGTHNATLQTSAGALVTMVT 462

QY 392 DDEKFFPMSGVM-----IFGKE-----SAGASNTALDNVMTDE 426
Db 463 NGAD---VSGVRAVRVGYSTDPYGGQQPESDLLRLRYSASAEQQNPILEN----- 512

QY 427 EEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVYLQGPWAKIP 486
Db 513 -----AARHTFTREATKLITGSGADGDYKEWMLPNQMWD SAPISRYNPIWVKVP 564

QY 487 HTDGHFHPSPMLMGGFGLKHPPPQILIK--NTPVPANPPAEFSATKFASFITQYSTGOVS 544
Db 565 RVNRKTLTLDQGSIPMSHPPGTIFIKLARIPVPGND-----SFLNIYVTGOVSC 615

QY 545 EIEWELQKENSKRWNPEVQYTSNYAKSANVD-FTVDNNGLY 584
Db 616 EVVWEVEKRGTKNWRPEYMHs---ATNMSVDAYTINNAGVY 653

coat protein VP1 - parvovirus B19 (strain Au)
C;Species: parvovirus B19
A;Note: host Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A24299
R;Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A;Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated from
A;Reference number: A24299; MUID:86200451; PMID:3701931
A;Accession: A24299
A;Molecule type: DNA
A;Residues: 1-781 <SHA>
A;Cross-references: EMBL:M13178; NID:g333375; PIDN:AAA66867.1; PID:g333377
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match 15.2%; Score 495; DB 1; Length 781;
Best Local Similarity 27.3%; Pred. No. 5.1e-26;
Matches 148; Conservative 84; Mismatches 257; Indels 54; Gaps 16;

QY	51	LGEPPATPAAGPTTMAAGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITSTRTW	110
Db	213	LPEVPAYNASEKYPMSMTSVNSAE-ASTGAGGGGNSVKSMWSEGATFSANSVTCTFSRQF	271
QY	111	ALPTYNNHLYKQISSASTGASND-----NHYFGYSTPWGYFDNRFHCHFSRDRW	160
Db	272	LIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGYSTPWRYLDFNALNLFSPLEF	331
QY	161	QRLINNWFGRPKRLNFKLFNIQKVT--TNDGVTTIANNLTSTVQFSDSEYQLPYVL	218
Db	332	QHLIENYGSIAPDALTVTTISEIAVKQVDTKGGV-QVDTSTTGRCLMLVDHEYKYPYVL	390
QY	219	GSAHQGCLPPFPADVFMIPQGYLTLNN-GSQAVG-----RSSFYCLEYFSPQMLR	268
Db	391	GQGQDTLAPELPIWYFPPQYAYLTVGDVNTQIGSGDSKKLASESAFYVLEHSSFQLLG	450
QY	269	TGNNFTFSYTFEEVPPFHSSYAHSQSLDRLMNLIDQYLYLNRNTQNSGSAQNKDLLFSR	328
Db	451	TGGTASMSYKFPFPPVPENLEGCSQHFEYEMYNPL---YGSRLGVDPDTLGGDPKFRSL	503
QY	329	GSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPG-TAM	387
Db	504	-THEDHAIQPNFPGPLVNSVSTKEGDSSTNGAGKALTGLSTGTSQNTTRISLRPGVVSQ	562
QY	388	ASHKDEDEKFFPMGSGVMIFGKESAGASNTALDNV-----MITDEEIKATNPVATERF	440
Db	563	PYHHWTDKVTGTGINAISHGQTTYG---NAEDKEYQQGVGRFPNEKEQLKQLGLNMHTY	619
QY	441	GTVAVNFQSSSTDPATGVDVHAMGALPGMVWQDRDVLQGP1WAKIPHDTGHFHPS-PLMG	499
Db	620	-----FPNKGTOQYTDQIE-RPLMVGSVMNRRAHYESQLWSKIPNLDDDSFKTQFAALG	672
QY	500	GFGLNKPPPPQILIKNTVPVPANPPAEFSATKPFASFITQYSTQGVSVIEWEL-QKENSKRW	558
Db	673	GWGLHQPPPPQIFLK--ILPQSGPIGGIKSMGITTLLVQYAVGIMVTMTVTFKLGPRKATGRW	730
QY	559	NPE 561	
Db	731	NPQ 733	

RESULT 5
VCPVPP
coat protein VP1 - porcine parvovirus
N;Contains: coat protein VP2
C;Species: porcine parvovirus
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 12-Apr-1996
C;Accession: B33302
R;Ranz, A.I.; Mancus, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A;Title: Porcine parvovirus: DNA sequence and genome organization.
A;Reference number: A33302; MUID:90010964; PMID:2794971
A;Accession: B33302
A;Molecule type: DNA
A;Residues: 1-723 <RAN>
A;Cross-references: EMBL:D00623
C;Genetics:

A;Introns: 10/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 8.0%; Score 258.5; DB 1; Length 723;
Best Local Similarity 23.1%; Pred. No. 7.8e-10;
Matches 155; Conservative 92; Mismatches 270; Indels 153; Gaps 35;

QY 9 VEQSPQEPDSSSGIGK-----TGQOPAKRLNFGQTGDSSEVPDPQPLGEPPTATP 58
Db 103 VRRSPRKHPGSKPPGKRPAPRHIFINLAKKAKGTSTNSNSMSENVEQHNPIN--AATE 160

QY 59 -AAVGPTTMAAGGAPMADNNEGADGVGNASGNWH--CDSTWLGD---RVITTSTRTWAL 112
Db 161 LSATGNESGGGGGG---GGRGAGGVGVSTGSFNNQTEFOYLGELGLVRITAHASRLIHL 216

QY 113 PTYNNHLYKQI-----SSASTGAS-NDNHYFGYSTPWGYFDNRFHCHFSRDRWQRLINN 166
Db 217 NMPEHETKYRIHVLNSESAGQMVQDDAHTQMTVPWSLIDANAWGVWFNPADWQLISNN 276

QY 167 NWGFRPKRLNFKLFNIQKVT---TNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQ 223
Db 277 MTEINLVSFEQAI FNVVLKTTITESATSPPTKIYNNDLTASLVALDTNNTLPTTPAAPRS 336

QY 224 GCLPFPFADVFMIPOGYL-----TLNNGSQAVGRSS-----FYCLE-YFP 263
Db 337 ETLGFYFWLPTKPTQYRYLYSCIRNLNPPYTYTGOSQPNRLNRLNRLHSDIMFYTIENAVP 396

QY 264 SQMLRTGNPF-TFSYTFEEVPPFHSSYAHSQSLDR-----LMNPLI--DQYLYLNR 312
Db 397 IHLRTGDEFSTGIYHFDTKPL--KLTHSWQTNRSRLGLPPLKLTETPTTEGDOHPGTLPA 454

QY 313 QNQSAGAKNDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKY 372
Db 455 NTRKGYHTINNSYTEAT-----AIRP-----AQVGYNTPYMNFYSNGGPF 496

QY 373 NLNGRESIINP--GTAMASHKDDDEDKFFPMGSGVMIFGKESAGASNTALD---NVMITDEE 427
Db 497 -----LTPIVPTADTQYNDDEPN-----GAIRFTMDYQHGHLLTSSQ 533

QY 428 EIK--ATNP-----VATERFCTVA-VNFQSSS-----TDPATG--DVHAMGALP---- 466
Db 534 ELERYTFNPQSKCGRAPKQFQQAAPLNLENTNNGTLTLPSPDFIGGKSNMHFMNTLNTYGP 593

QY 467 -GMVWQDRDVLQGP1WAKIPHDT--GHFHPSPLMGFGGLK-NPPPPQILIKNTVPVPANPP 522
Db 594 LTALNNTAPVPFNGQIWDKELDTDLKPRLH---VTAPFVCKNPPPGQLFVKIAP---NLT 647

QY 523 AEESA-TKFASFITQYSTQGVSVIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNN 581
Db 648 DDFNADSPQQPRIITYSNFWWKGTLTFTAKMRSSNMWNP1QQHTT-----TAENI 697

QY 582 GLYTEPRRPIG 591
Db 698 GKVI-PTNIG 706

RESULT 6
VCPVNA
coat protein VP1 - porcine parvovirus (strain NADL-2)
N;Contains: coat protein VP2
C;Species: porcine parvovirus
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: B33743; D48472
R;Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 173, 368-377, 1989
A;Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pa
A;Reference number: A33743; MUID:90085785; PMID:2596019
A;Accession: B33743
A;Molecule type: DNA
A;Residues: 1-729 <VAS>
A;Cross-references: GB:M32787; NID:g3332983; PIDN:AAA46917.1; PID:g3332985

C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Apr-1994
C;Accession: A03699
R;Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization
A;Reference number: A03695; MUID:83112183; PMID:6823009
A;Accession: A03699
A;Molecule type: DNA
A;Residues: 1-722 <RHO>
A;Cross-references: EMBL:X01457; EMBL:J02198
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match	7.5%;	Score 244;	DB 1;	Length 722;
Best Local Similarity	21.4%;	Pred. No. 7.7e-09;		
Matches	142;	Conservative 94;	Mismatches 267;	Indels 160; Gaps 29;

Qy 8 PVEQSPQEPDSSGIGKTGQ---PA-----KKRLNFGQTGDSESVDPQPPLGEP 54
Db 82 PKLSTDSEP-GTSGVSRPGKRTKPPAHIFVNQARAKKRASLAQAQRTLTMSDGTETNQP 140

Qy 55 PATPAAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTW--LGDRVITSTRTWAL 112
Db 141 DTGIANARVERSADGGGS-SGGGGSGGGIGVSTGTVDNQTTTKFLGDGWVEITAHASRL 199

Qy 113 -----PTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSR 158
Db 200 LHLGMPPESENCRVTVHNNQTTGHGTVKVGKNMAYDTHQIIV-TPWSLVDANAWGVWFQPS 258

Qy 159 DWORLINNWGFRPKRLNFKLFNIQVKEVTT-----NDGVTTIANNLTSTVQVFSSEYQ 213
Db 259 DWQFIONMESLNDLSLQELFNVVVKTVEQQGAGQDAIKVYNNDLTACMVVALDSNNI 318

Qy 214 LPYVLGSAHQCL-----PPFPAD---VFMIPOGYLTLNNGSQ-----AVG-----R 253
Db 319 LPYTPAAQTSFGLGYFPWKPTAPAPYRYFFMPRQLSSTSSAEGTQITDTIGEPQALN 378

Qy 254 SSFYCLE-YFPSQMLRTGNFTF-SYTFEEVPFHSSYA-----HSQSLDRLMN-PLIDQY 305
Db 379 SQFPTIENTLPIITLLRTGDEFTTGTIYFNTDPLKLTHTWTQNRHLACLQGITDLPTSDTA 438

Qy 306 LYLYNRTQNQSGSAQNKDLF-----SRGSPAGMSVQPKNWL 342
Db 439 TASLTANGDRFGSTQTQNVNVVTEALRTPAQIGFMQPHDNFENRGFPFKVPVVP----- 494

Qy 343 PGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGR--ESIINPGTA-----MASH 390
Db 495 -----LDITAGEDHDAN---GAIRFNYGKHGEDWAKQGAAPERYTWDIDAISAG 541

Qy 391 KDEDKFFPMMSGVMIFGKESAGASNTALDNVMTDDEEIIKA-TNPVATERFGTVAVNFOS 449
Db 542 RDTARCFV-----QSAPISIPENQOILQREDIAIGRTNMHYTNVFNFSYGPLSAF 591

Qy 450 SSTDPATGDVHAMGALPGMVWQDRDVLQGPWAK---IPHTDGHFHPSPMLGGFGLKNP 506
Db 592 PHPDP-----IYPNGQIWDKELDLHKPRLLHVTAPFV---CKKNP 628

Qy 507 PPQILIKNTPVPANPPAEF--SATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQY 564
Db 629 PGQLFVHLGP---NLTDQFDPNSTTVSRIVT-YSTFYWKGLKFKAKLRPNLTWNPVYQA 684

Qy 565 TSN 567
Db 685 TTD 687

RESULT 9
B44276
coat protein VP1 - parvovirus LuIII
C;Species: parvovirus LuIII
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C;Accession: B44276
R;Diffot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.

Virology 192, 339-345, 1993
A;Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A;Reference number: A44276; MUID:93297126; PMID:8517025
A;Accession: B44276
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-587 <DIF>
A;Cross-references: GB:M81888
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein; glycoprotein
F;49,90,220,304,371,503,511,514,539,571/Binding site: carbohydrate (asn) (covalent) #sta

Query Match	7.2%;	Score 234;	DB 1;	Length 587;
Best Local Similarity	23.1%;	Pred. No. 2.8e-08;		
Matches	137;	Conservative 74;	Mismatches 231;	Indels 152; Gaps 29;

Qy 59 AAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCHDS--TWLGD---RVITTSTRTWAL- 112
Db 23 AADGPG--GSGGG-----GSGGGGVGVSTGSDNQTHYKFLGDGWVEITAYSTRMVHLN 74

Qy 113 -PTYNN-----HLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWORLIN 165
Db 75 MPKSENYCRVRVHNTNDGTASHMAMDDAHEQIW-TPWSLVDANAWGVWFQPSDWQYISN 133

Qy 166 NNWGFPRKRLNFKLFNIQVKEVT-TNDGVTTTIA---NNLTSTVQVFSSEYQLPYVLGSA 221
Db 134 NMTHINLSLDQELFNVIKTVTEQNTGAEAIKVYNNDLTAAAMVALDSNNILPYTPAID 193

Qy 222 HQCL-----PPFPA-----DVFMIPQGYLTLNNGSQAVGRSSFYCLEYFP 263
Db 194 NQETLGFYPWKPTIPSPYRYFFSCDRNLSVTYKDEAGTITDTMGLASGLNSQFFTIENQ 253

Qy 264 S-QMLRTGNF-TFSYTFEEVPFHSSYAHSSQSLDRLMNLIDQ-----YL 306
Db 254 RINLLRTGDEYATGTYFDTEPIRLTHTWTQNRHLGQPPQITELPSSDTANATLTARGYR 313

Qy 307 YLYNRTQN-----QSGSAQNKD-LLFSRGPAGMSVQPKNWLPG----- 344
Db 314 SGLTQIQGRNDVTEATVRPAQVGFQCPHDNFETSRAGPFKVPVVPADITQGLDHDANGS 373

Qy 345 --PCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESINPGTAMASHKODEDKFFPMMSG 402
Db 374 LRYTYDKQHGQSWASQNNKDRYTW-DAVNYD-SGR----- 406

Qy 403 VMIFGESAGASNTALDNVMTDDEEIIKATNPVATER---FGTVAVNFQS--SSTDPA TG 457
Db 407 -----WTNNCFIQSVFPTSEPN---ANQILTRNDNLAKGTDIHTFNAFNSYGLPLTA 454

Qy 458 DVHAMGALPGMVWQDRDVLQGPWAK---IPHTDGHFHPSPMLGGFGLKNPPQILIKN 514
Db 455 FPH-----PAPIYQGGQIWDKELDLHKPRLLHTQAPFV---CKKNAPGQLLVRL 500

Qy 515 TP--VPANPPAEFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEVQYTS 566
Db 501 APNLTDQYDPNSSLRSRIVTYGTFFWKGLTLKAK---MRPNA-TWNPVFQISA 550

RESULT 10
VCPV2M
coat protein VP1 - minute virus of mice
C;Species: minute virus of mice, murine parvovirus
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Apr-1994
C;Accession: A03700
R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.
Nucleic Acids Res. 11, 999-1018, 1983
A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.
A;Reference number: A03696; MUID:83143341; PMID:6298737
A;Accession: A03700
A;Molecule type: DNA
A;Residues: 1-716 <AST>
A;Cross-references: EMBL:V01115
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match	6.6%;	Score 216;	DB 1;	Length 716;
Best Local Similarity	20.8%;	Pred. NO. 6.3e-07;		
Matches 148;	Conservative 93;	Mismatches 271;	Indels 198;	Gaps 34;

QY	8	PVEQSPQEPDSSSGIGKTG-----QQPAKURL-----NFGQT-GDSESVDPD	48
Db	82	PKLATDSEP-GTSGVSRAGKTRPPAYIFINQARAKKLTSSAAQQSSQTSMDGTSQPD	140
QY	49	QPLGEPPA--TPAAVGPTTMASSGGGAPMADNNEGADVGNASGNWHCHDS--TWLGD---	101
Db	141	GNAVHSAARVERAADGPG--GSGGG-----GSGGGGVGVSTGSYDNQTHYRFLGDGW	192
QY	102	VITSTRTWALPTYNNHLYKQI-----SSASTGASNDNHYFGYSTPWGYPDFNRFCH	154
Db	193	ITALATRLVHLNMPKSENYCRIRVHNTTDSVKGNMAKDDAHEQIWTPWSLVDANAWGV	252
QY	155	FSPRDWQRLLNNWGFPRKRLNFKLFNIQVKEVTND---GVTTIANNLSTVQVFSDS	210
Db	253	LQPSDWQYICNTMSQLNLVSLDQEIFNVVLKTVTEQDLGGQAIKIYNNDLTACMVA	312
QY	211	EYQLPYVLGSAHQGLPPFPADVFMIPOGY-----LTLNNGSQAV-----G	252
Db	313	NNILPYTPAANSMETLGFYPWKPTIASPYRYFCVDRDLSVTYENQEGTVEHNVMTGPK	372
QY	253	RSSFYCLEYFPS-QMLRTGNF-TFSYTFEEVPFHSSYAHQSOLDRLMN--PLIDQYLY	308
Db	373	IPQFFTIENTQITLLRTGDEFATGTYYFDTSV--KLTHWTQNRQLGQPPLSTF---	427
QY	309	LNRTQSGSAQNKDLLFSRGSPAGMSVQPKNWL-----PG	344
Db	428	PEADTDAGT-----LTAQGRHGTTQMGVNVWVSEAIRTPAQVGFQCPHNDFEASRAG	480
QY	345	PCYRQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPG-----TAMASHKODE	394
Db	481	P-FAAPKVPADITQGVNKEANGSVRYSGKHGENWASHGPAPERYTWDETSPGSGRDTK	539
QY	395	DKFF-----PMSGVMI FGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAV	445
Db	540	DGFIQASAPLVVPPPLNGI-----LTNANPFGTKN-----DI	570
QY	446	NFQS--SSTD PATGDVHAMGALPGMVWQDRDYVLQGPWAK---IPHTDGHFHPSPLMGG	500
Db	571	HFSNVFNSYGLTAPSH-----PSPVYFQGIWDKELDLHKPRLHITAPFV--	617
QY	501	FGLKNPPPPQILLKNTP--VPANPPAEFSATKEASFITQYSTGQVSVEIEWELQENSKRW	558
Db	618	-CKNNAPGQMLVRLGPNLTDQYDPNGATLSRIVTYGTFFWKGLTMRACLRA-----NTTW	672
QY	559	NPEVQYTSNYAKSANVDFTVDNNGLY---TEPRPIGT-----RYLTRPL	599
Db	673	NPVYQVSAE-----DNGNSYMSVTXWLPATATGNMQSVPLITRPV	711

RESULT 11
VCPVIF
coat protein VP1 - feline panleukopenia virus
N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03701
R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Hahn, W.
J. Virol. 55, 574-587, 1985
A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A;Reference number: A03697; MUID:85265017; PMID:2991581
A;Accession: A03701
A;Molecule type: DNA
A;Residues: 1-727 <CAR>
A;Cross-references: EMBL:M10824; NID:g3333474; PIDN:AAA47161.1; PID:g3333476
C;Genetics:
A;Introns: 11/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

F;144-727/Product: coat protein vp2 #status predicted <VP2>

Query Match 6.6%; Score 216; DB 1; Length 727;
Best Local Similarity 20.2%; Pred. NO. 6.5e-07;
Matches 139; Conservative 101; Mismatches 259; Indels 188; Gaps 34;

QY	7	RPVEQSPQEDSSSGIGKTGQQPAKRLNFGOTGDSVP----	DPQPLGEPPA--TPAA	60
DB	110	KPTKRSKPPPHIFINLAK-----KKKAGAGQVKRDNQAPMSDGA	VQPDGGQPAVRNERA	163
QY	61	VGPTTMAASGGGAPMADNNEGADGVGNASGNWH-----	CDSTWLGDRVITTTSTRTWALP	113
DB	164	TSGSGNGSGGG-----GGSGGVGISTGTFTNNQTEFKFLENGWV--	EITANSSRLVHLN	215
QY	114	TNNHLYKQI-----SSASTGASNDNHYFGYSTPWGYDFNRFHCH	FSRDPQWQLINN	166
DB	216	MPESENYKRVVNNMDKTAVKGNMALDDTHVQIVTPWSLVDANAWGV	WFENPGDWQLIVNT	275
QY	167	NWGFPRKRLNFKLFNIQKVEVT--TNDGVTTIANNLTTSTVQVFS	DSEYQLPYVLGSAHQ	223
DB	276	MSELHLVSEQEIFNVVLKTVSESATQPPTKVYNDLTASLMVALDS	NNTMPFTPAAMRS	335
QY	224	GCL-----PPPPA-----DVFMIQYGYLTLNNGSQAVGRS-	-----SFYCL	259
DB	336	ETLGFYPWKPTIPTPWRYFFQWDRRTLIPSH-----TGTS	GTPTNIYHGTDPPDVQFYTI	389
QY	260	E-YFPSQMLRTGNMF-TFSYTFEEVPPHSSVAHSQSLDRLLMN--	PLIDQYLYLNRTOQ	315
DB	390	ENSVPVHLRLTGDEFATGTFDFCKP--CRLTHWTQNRALGLPP-	-----FLNSLPQS	440
QY	316	SGSAQNKOLLFSRGSPPAGMS-----VQPKN-WLPGPCYRQORVSK-	-----	354
DB	441	EGATNFGDIGVQDDKRRGVQTQMGNTDYITEATIMRPAEVGYSAPY	SFEASTQGPFKIP	500
QY	355	-----TKTDNNSNFTWTGASKY--NLNGRESINPGT----	AMASHKODEDKFFPMS	401
DB	501	AAGRGAQTDENGAA--DGDPRYAFGRQHGGKTTTGETPERFTYIA	HQDT-----	549
QY	402	GVMIFGKESAGASNTALD-NVMITDEEIKATNPVATERFGTVAVNF	QSSSTDPATGDVH	460
DB	550	-----GRYPAGDWIQNINFNLPVTNDNVLLPTDPIG-----	GKTGINY--TNIFNTYGPLT	598
QY	461	AMGALPGMWQDRDVYLGQPIWAKIPHTDGHFHPSLMGGFGLK-	-----NPP	507
DB	599	ALNNVP-----PVYPNGQIWDKEFDTD-----	LKPRLHVNPFFVCQNNCP	638
QY	508	PQILIKNTVPVPA-----PPAEFSATKFASFITQYSTQVSV	EIEWELQENSKRWNP	564
DB	639	QGLFVKVAPNLTNEDYDPDASANMSR-----IVTYSDF	WWKGLVFKAKLRASHTWNPIQ	694
QY	565	TSNYAKSANVDFTVDNNGLYTEPRPIG	591	
DB	695	SIN-----VDNQFNYL-PNNIG	710	

RESULT 12

VCPVIM

coat protein VP1 - minute virus of mice (strain MVMI)

C:Species: minute virus of mice, murine parvovirus

C:\projects\mmsc\src\mmsc\main.f90

C>Date:	28-Dec-1987	#sequence	revision	28-Dec-1987	#text	change
						28-Jul-2000

C:Accession: B23008: B29510

R: Sahli. B.: McMaster. G.K.: Hirt. B.:
C, ACCESION: D23000, D23010

Nucleic Acids Res. 13: 3617-3633, 1985

A-Title: DNA sequence comparison between two tissue-specific variants of the autonomous NUCLEIC ACIDS RES. 13, 3617-3633, 1983

A; title: DNA sequence comparison between two CIB8de-8pe
A: reference number: A23008: MUID: 85242059: PMID: 3855242

A;Reference Number: A23
A:Accession: B23008

A; Accession: B23008
A: Molecular mass: DNA

A; Molecule type: DNA
A: Position: 1 718 - 5848

A;Residues: 1-718 <SAH>
 NCBI accession: EMBL:Y02481. NID:750010. PDB:CAP45E07 1. PDB:5E410928

A;Cross-references: EMBL:X02481; NID:960918; E

R;Astell, C.R.; Gardiner, E

J. Virol. 57, 656-669, 1986

A;Title: DNA sequence of the lymphotropic variant of mi

A;Accession: B29510
A;Molecule type: DNA
A;Residues: 1-143,'A',145-718 <AST>
A;Cross-references: EMBL:M12032
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match 6.6%; Score 214.5; DB 1; Length 718;
Best Local Similarity 21.0%; Pred. No. 8.1e-07;
Matches 142; Conservative 89; Mismatches 273; Indels 171; Gaps 30;

```
QY      8 PVEQSPQEPDSSSGIGKGTG-----QOPAKRLNFG-----QTGDSESVDPQPLG 52
Db      83 PKLATDSEP-GTSGVSRAGKRTRPPAYIFINQARAKKLTSSAAQSSQTMSDGTSPDG 141

QY      53 -----EPPATPAAGVPTTMASGGGAPMADNNEGADVGNASGNWHCDS--TWLGD---R 101
Db      142 GNGVHSAARVERAADGPG--GSGGG-----GSGGGGVGVSTGSYDNQTHYRFLGDGWE 193

QY      102 VITSTRITWALPTYNNHLYKQI-----SSASTGASNDNHYFGYSTPWPGYFDENRFCH 154
Db      194 ITALATRLVHLNMPKSENYCRIRVHNTTDTSVKGNMAKDDAHEQIWTWPWSLVDANAWGV 253

QY      155 FSPRDQRLINNNWGFRLKLNFKLNIQVKEVTND---GVTTIANNLTSTVQVFSDS 210
Db      254 LQPSDWQYICNTMSQLNLVSLDQEIFNVVLKTVTEQDSGGGQAIKYNDLTACMMVAVDS 313

QY      211 EYQLPYVLGSAHQGLPPFPADVFMIPOYGY-----LTLNN-----GSOAV 251
Db      314 NNILPYTPAANSMETLGFYPWKPTIASPYRYFCVDRDLSVTYENQEGTIEHNVMGTPKG 373

QY      252 GRSSFYCLEYFPPS-QMLRTGNF-TFSYTFEEVFPFHSSYAHQSQSLDRLMN--PLIDQVLY 307
Db      374 MNSQFFTIENQOITLLRTGDEFATGYTFDTNPV--KLTHWTQNRQLGQPPLLSTF-- 429

QY      308 YLNRTQNSGSAQNKKOLLFSRGSFAGMSVQPKNWL-----P 343
Db      430 --PEADTDAGT-----LTAQGSRHGATQMEVNVWVSEAIRTPAQVGFQCPHNDFEASRA 481

QY      344 GPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDF----- 397
Db      482 GP-FAAPKVPADVTQGVDRANGSVRSYSGKHGENWAAHGPAPERYTWDETNGSGRDT 540

QY      398 ---FPMGVMIFGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQS--SST 452
Db      541 RDGFIQSAPLVVPVPPPLNGILTNA-----NPIGTKN----DIHFSNVFNSY 581

QY      453 DPATGDVHAMGALPGMVWQDRDVYLQGPWIAK---IPHTDGHFSPPLMGFGGLKNPPBQ 509
Db      582 GPLTAFSH-----PSPVYPQGIWDKELDLHKPLRHITAPFV---CKNNAPCQ 627

QY      510 ILIKNTP--VPANPPAEFSATKFAFITQYSTGQVSVEIEWELQKENSKRWNPEVQTSN 567
Db      628 MLVRLGPNLTDQYDPNGATLSRIVTYGTFFWKGKLTMRACLRA-----NTTNPVYQ---- 679

QY      568 YAKSANVDFTVDNNG 582
Db      680 -----VSVEDNG 686
```

RESULT 13
VCPVMB
coat protein VP1 - mink enteritis virus (strain Abashiri)
N;Contains: coat protein VP2
C;Species: mink enteritis virus, MEV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Apr-1996
C;Accession: B38350
R;Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishigurio, N.; Goto, H.; Shinagawa, J. Gen. Virol. 72, 867-875, 1991
A;Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: B38350
A;Molecule type: DNA

A;Residues: 1-722 <KAR>
A;Cross-references: GB:D00765
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;139-722/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.6%; Score 214; DB 1; Length 722;
Best Local Similarity 20.2%; Pred. No. 8.8e-07;
Matches 139; Conservative 101; Mismatches 259; Indels 188; Gaps 34;

```
QY      7 RPVEQSPQEPDSSSGIGKGTQQQPAKRLNFGQTGDSESV-----DPOPLGEPPA--TPAA 60
Db      105 KPTKRSKPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAVOPDGGQPAVRNERA 158

QY      61 VQPTTMASGGGAPMADNNEGADVGNASGNWH-----CDSTWLGDRVITSTRTWALP 113
Db      159 TSGSGNGSGGG-----GGSGGVGISTGTFFNNQTEFKPLENGWV--EITANSSRLVHLN 210

QY      114 TYNNHLYKQI-----SSASTGASNDNHYFGYSTPWPGYFDENRFCHFSRDRWQRLINN 166
Db      211 MPESENYKRVVNVNNDKTAVKGNMALDDTHVQIVTPWSLVDANAWGVFNPGDWQLIVNT 270

QY      167 NWGFRPKRLNFKLNIQVKEVT--TNDGVTTIANNLTSTVQVFSQSEYQLPYVLGSAHQ 223
Db      271 MSELHLVSFEQEIFNVVLKTVSESATQPTTKVYNNNDLTASLMVALDSNNTMPPTPAAMRS 330

QY      224 GCL-----PPFPA-----DVFMIPQGYLTLNNGSQAVGRS-----SFYCL 259
Db      331 ETLGFYPWKPTIPTPWRYYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDDVQPYTI 384

QY      260 E-YFPSQMLRTGNF-TFSYTFEEVFPFHSSYAHQSQSLDRLMN--PLIDQYLYLNRQNTQ 315
Db      385 ENSVPVHLLRTGDEFATGTFEFDCKP--CRLTHWTQNRALGLPP-----FLNSLPQS 435

QY      316 SGSANQKDLLFSRGSFAGMS-----VQPKN-WLPGPCYRQQRVSK----- 354
Db      436 EGATNFGDIGVQDKRQGVQMGNTDYITEATIMRPAEVGYSAFYYSFEASTQGPPEKPTPI 495

QY      355 -----TKTDNNNSNFTWTGASKY---NLNGRESIINPGT---AMASHKDDKDFPMS 401
Db      496 AAGRGAQTDENQAA---DGDPRYAFGRQHGQKTTTGETPERFTYIAHQDT----- 544

QY      402 GVMIFGKESAGASNTALD-NVMTDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVH 460
Db      545 -----GRYPAGDWIQNINFNLPVTNDNVLLPTDPIG---GKTGINY--TNIFNTYGPLT 593

QY      461 AMGALPGMVWQDRDVYLQGPWIAKIPHTDGHFSPPLMGFGGLK-----NPP 507
Db      594 ALNNVP-----PVYPNGQIWDKEFDTD-----LKPRLHVNPAPFVCQNNCP 633

QY      508 PQILIKNTPVAN---PPAEFSATKFAFITQYSTGQVSVEIEWELQKENSKRWNPEVQY 564
Db      634 GQLFVKVAPNLTNEDPDASANMSR-----IVTYSDFWKKGLVFKAKLRASHTWNPQQM 689

QY      565 TSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db      690 SIN-----VDNQFNYL-PNNIG 705
```

RESULT 14
VCPVFP
coat protein VP1 - feline panleukopenia virus (strain 193)
N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 12-Apr-1996
C;Accession: B36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvo A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: B36608
A;Molecule type: DNA
A;Residues: 1-727 <MAR>

A;Cross-references: GB:X55115
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.4%; Score 209; DB 1; Length 727;
Best Local Similarity 20.1%; Pred. No. 2e-06;
Matches 138; Conservative 101; Mismatches 260; Indels 188; Gaps 34;

QY 7 RPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTDSESV-----DPQPLGEPPA--TPAA 60
Db 110 KPTKRSKPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAQVQDGGQPAVRNERA 163
QY 61 VGPTTASGGGAPMADNNEGADGVGNASGNWH-----CDSTWLGDRTVITSTRTWALP 113
Db 164 TSGNGSGGGG-----GGSGGVGISTGTFFNQTQEFKFLNGWV--EITANSSRLVHLN 215
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPWGYFDENRFCHFSRDRWQRLINN 166
Db 216 MPESYKRVVNNMDKTAVKGNMALDDIHVQIVTPWSLVDANAWGVWENPGDWQLIVNT 275
QY 167 NWGFRPKRLNFKLNIQKVEVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQ 223
Db 276 MSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLVALDSNNTMFTPAAMRS 335
QY 224 GCL-----PPFPA-----DVFMIPQYGLTLNNGSQAVGRS-----SFYCL 259
Db 336 ETLGFYPWKPTIPTPRYFYQWDRTLIPSH-----TGTSGTPTNVYHGTDPDDVQFYTI 389
QY 260 E-YFPSQMLRTGNF-TFSYTFEEVFPFHSSYAHQSOLDRLMN--PLIDQYLYLNRTQNO 315
Db 390 ENSVPVHLRLTGDEFATGTFDFCKP--CRLTHWTQTNRALGLPP-----FLNSLPQS 440
QY 316 SGSAQNKDLLFSRGSPPAGMS-----VQPKN-WLPGCYRQQRVSK----- 354
Db 441 EGATNFGDIGVQDQKRRGVYQMGNTDYITEATIMRPAEVGYSAPIYSFEASTQGPFTPI 500
QY 355 -----TKTDNNNSNFTWTGASKY---NLNGRESINPGT---AMASHKDDDKFFPMS 401
Db 501 AAGRGAQTDENQAA---DGDPRYAFGRQHGQKTTTGETPERFTYIAHQDT----- 549
QY 402 GVMIFGKESAGASNTALD-NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPAVDVH 460
Db 550 -----GRYPEGDWIONINFLPVTNDNVLLPTDPIG-----GKTGINY--TNIFNYGPLT 598
QY 461 AMGALPGMVWQDRDVLQGIWAKIPHTDGHFHPSPLMGGFGLK-----NPP 507
Db 599 ALNNVP-----PVYPNGQIWDKEFDDT-----LKPRLHVNAPEVCQNNCP 638
QY 508 PQILIKNTVPVAN---PPAEFSATKFPASFITQYSTGQVSVEIEWELQKENSKRWPEVQY 564
Db 639 GOLPVKVPAPNLITNEYDPDASANMSR----IVTYSDFWFKGLVFKAKLRASHTWNPFIQOM 694
QY 565 TSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 695 SIN-----VDNQFNIV-PNNIG 710

RESULT 15
VCPVCD
coat protein VP1 - canine parvovirus (strain CPV-d)
N;Contains: coat protein VP2
C;Species: canine parvovirus, CPV
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A31163
R;Parrish, C.R.; Aquadro, C.F.; Carmichael, L.E.
Virology 166, 293-307, 1988
A;Title: Canine host range and a specific epitope map along with variant sequences in th
A;Reference number: A31163; MUID:89020796; PMID:3176341
A;Accession: A31163
A;Molecule type: DNA
A;Residues: 1-737 <PAR>
A;Cross-references: EMBL:M23255; NID:G3333467; PIDN:AAA47158.1; PID:G3333468

C;Genetics:
A;Introns: 26/3
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.4%; Score 208; DB 1; Length 737;
Best Local Similarity 19.8%; Pred. No. 2.3e-06;
Matches 137; Conservative 96; Mismatches 260; Indels 200; Gaps 31;

QY 7 RPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTDSESV-----DPQPLGEPPA--TPAA 60
Db 120 KPTKRSKPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAQVQDGGQPAVRNERA 173
QY 61 VGPTTASGGGAPMADNNEGADGVGNASGNWH-----CDSTWLGDRTVITSTRTWALP 113
Db 174 TSGNGSGGGG-----GGSGGVGISTGTFFNQTQEFKFLNGWV--EITANSSRLVHLN 225
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPWGYFDENRFCHFSRDRWQRLINN 166
Db 226 MPESYKRVVNNMDKTAVKGNMALDDIHVQIVTPWSLVDANAWGVWENPGDWQLIVNT 285
QY 167 NWGFRPKRLNFKLNIQKVEVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQ 223
Db 286 MSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLVALDSNNTMFTPAAMRS 345
QY 224 GCL-----PPFPA-----DVFMIPQYGLTLNNGSQAVGRS-----SFYCL 259
Db 346 ETLGFYPWKPTIPTPRYFYQWDRTLIPSH-----TGTSGTPTNVYHGTDPDDVQFYTI 399
QY 260 E-YFPSQMLRTGNF-TFSYTFEEVFPFHSSYAHQSOLDRLMN--PLIDQYLYLNRTQNO 315
Db 400 ENSVPVHLRLTGDEFATGTFDFCKP--CRLTHWTQTNRALGLPP-----FLNSLPQS 450
QY 316 SGSAQNKDLLFSRGSPPAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSNF-----TW 366
Db 451 EGATNFGDI-----GV-----QDKRRGVYQMGNTNYITEATIMRPAE 488
QY 367 TGASKYNLNGRESIINP-----GTAMASHKDDDKFFPMSGVMIFGKESAGASNTALD-- 419
Db 489 VGYSAPYYSFEASTQGPFTPIAAGRGAQTDENQAAADGNPRYAFGRQHGQKTTTGETP 548
QY 420 -----NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPA 454
Db 549 ERFTYIAHQDTGRYPEGDWIONINFLPVTNDNVLLPTDPIG-----GKTGINY--TNIFN 602
QY 455 ATGDVHAMGALPGMVWQDRDVLQGIWAKIPHTDGHFHPSPLMGGFGLK----- 504
Db 603 TYGLPLTALNNVP-----PVYPNGQIWDKEFDDT-----LKPRLHVNAPEV 642
QY 505 ---NPPQILIKNTVPVAN---PPAEFSATKFPASFITQYSTGQVSVEIEWELQKENSKRW 558
Db 643 QONNCPGQLFVKVAPNLITNEYDPDASANMSR----IVTYSDFWFKGLVFKAKLRASHTW 698
QY 559 NPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 699 NPIQQMSIN-----VDNQFNIV-PNNIG 720

Search completed: January 21, 2004, 16:07:07
Job time : 16.1017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 21, 2004, 15:59:46 ; Search time 10.2558 Seconds
(without alignments)
2746.653 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2177.5	67.0	504	COA3_AAV2	P03135 adeno-assoc
2	556	17.1	673	COAT_PAVBO	P07297 bovine parv
3	495	15.2	781	COAT_PAVHB	P07299 human parvo
4	254.5	7.8	729	COA1_PAVP9	P33484 porcine par
5	254.5	7.8	729	COA1_PAVPN	P18546 porcine par
6	252.5	7.8	729	COA1_PAVPK	P52501 porcine par
7	244	7.5	722	COAT_PAVHH	P03136 hamster par
8	234	7.2	587	COAT_PAVL3	P36310 parvovirus
9	216	6.6	716	COAT_MUMIV	P03137 murine minu
10	216	6.6	727	COAT_FPV	P04864 feline panl
11	214.5	6.6	718	COAT_MUMIM	P07302 murine minu
12	214	6.6	722	COAT_MEVA	P27437 mink enteri
13	209	6.4	727	COAT_FPV19	P24840 feline panl
14	208	6.4	727	COAT_PAVCB	Q11213 canine parv
15	208	6.4	737	COAT_PAVCD	P17455 canine parv
16	204	6.3	748	COAT_PAVCN	P12930 canine parv
17	195	6.0	584	COAT_PAVC2	P30129 canine parv
18	193	5.9	722	COAT_PAVC7	P04863 canine parv
19	186.5	5.7	647	COAT_ADVG	P24029 aleutian mi
20	143.5	4.4	648	FXN1_MOUSE	Q61575 mus musculu
21	123	3.8	648	FXN1_HUMAN	O15353 homo sapien
22	123	3.8	1142	ENAM_PIG	O97939 sus scrofa
23	121	3.7	880	SYV_BACST	P11931 bacillus st
24	121	3.7	1113	N116_YEAST	Q02630 saccharomyc
25	119	3.7	667	ITF2_HUMAN	P15884 homo sapien
26	116.5	3.6	642	ITF2_CANFA	P15881 canis famil
27	113	3.5	781	SP3_HUMAN	Q02447 homo sapien
28	110.5	3.4	655	CDGT_KLEPN	P08704 klebsiella
29	110.5	3.4	1742	GUNA_CALSA	P22534 caldocellum
30	109.5	3.4	531	HEXB_PIG	Q29548 sus scrofa
31	109.5	3.4	670	ITF2_MOUSE	Q60722 mus musculu
32	109	3.4	1186	CEAA_BACTS	Q45710 bacillus th
33	108.5	3.3	860	AREA_PENRO	O13508 penicillium

34	108.5	3.3	1379	1	YFF9_SCHPO	O14066 schizosacch
35	108.5	3.3	1849	1	IGA4_HAEIN	P45386 haemophilus
36	107	3.3	2493	1	CYAA_USTMA	P49606 ustilago ma
37	106.5	3.3	1070	1	Y355_HUMAN	O15063 homo sapien
38	106	3.3	1271	1	Y338_MYCGE	P47580 mycoplasma
39	105.5	3.2	559	1	HNPB_PIG	Q03365 sus scrofa
40	105	3.2	1396	1	ITA2_DROME	P12080 drosophila
41	105	3.2	1656	1	OMP8_RICJA	O06653 r outer mem
42	104.5	3.2	717	1	DXS_ARATH	Q38854 arabidopsis
43	104.5	3.2	1122	1	ADP1_MYCGA	Q49379 mycoplasma
44	104	3.2	395	1	MAF1_YEAST	P41910 saccharomyc
45	103.5	3.2	1196	1	AMYB_PAEPO	P21543 paenibacill

ALIGNMENTS

RESULT 1
COA3_AAV2
ID COA3_AAV2 STANDARD; PRT; 504 AA.
AC P03135;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Probable coat protein 3.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164299; PubMed=6300419;
RA Srivastava A., Lusby E.W., Berns K.I.;
RT "Nucleotide sequence and organization of the adeno-associated virus 2 genome."
RL J. Virol. 45:555-564 (1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J01901; AAA42376.1; -.
DR PIR; A03698; VCPV3A.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
SQ SEQUENCE 504 AA; 56366 MW; 75899B017052B6B2 CRC64;

Query Match 67.0%; Score 2177.5; DB 1; Length 504;
Best Local Similarity 82.5%; Pred. No. 8.1e-137;
Matches 392; Conservative 33; Mismatches 47; Indels 3; Gaps 2;

Qy	66	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDVRVTTSTRTWALPTYNHLYKQISS	125
Db	1	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITSTRTWALPTYNHLYKQISS	60
Qy	126	ASTGASNDNHFGYSTPWGYFDNRFHCHFSRPDQRLNNNWGFRPKRLNFKLFNIQVK	185
Db	61	QS-GASNDNHFGYSTPWGYFDNRFHCHFSRPDQRLNNNWGFRPKRLNFKLFNIQVK	119
Qy	186	EVTNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQCLPPPPADVFEMIPOYGYLTILN	245
Db	120	EVTQNDGTTTIANNLTSTVQFTDSEYQLPYVLGSAHQCLPPFPADVFMPVPOYGYLTILN	179
Qy	246	NGSQAVGRSSPYCLEYFSPQMLRTGNFTPSYTFEEVFHSSYAHQSQSLDRMLNPLIDQY	305
Db	180	NGSQAVGRSSPYCLEYFSPQMLRTGNFTPSYTFEDVPFHSSYAHQSQSLDRMLNPLIDQY	239
Qy	306	LYLNRNTQNSGSAQNKDLFSSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT	365

Db 240 LYLSTNTSPGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQORVSKTSADNNNSEYS 299
QY 366 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFFGKESAGASNTALDNVMTID 425
Db 300 WTGATKYHLNGRDSLVP--AMASHKODEEKKFFPQSGVLIFGKQSEKTNVNEKVMITD 357
QY 426 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWIWAKI 485
Db 358 EEEIGTTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMWQDRDVLQGPWIWAKI 417
QY 486 PHTDGHFHPSPMLMGFGFLKHPPPQILIKNTVPANPPAEFSATKFAFITQYSTG 540
Db 418 PHTDGHFHPSPMLMGFGFLKHPPPQILIKNTVPANPSTTFSAAKFAFITQYSTG 472

RESULT 2
COAT_PAVBO STANDARD; PRT; 673 AA.
AC P07297; Q84374;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP2 [Contains: Coat protein VP3].
OS Bovine parvovirus (BPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10784;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87061184; PubMed=3783814;
RA Chen K.C., Shull B.C., Moses E.A., Lederman M., Stout E.R.,
RA Bates R.C.;
RT "Complete nucleotide sequence and genome organization of bovine
parvovirus.";
RL J. Virol. 60:1085-1097(1986).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14363; AAB59848.1; -;
CC EMBL; M14363; AAB59849.1; -;
CC PIR; A26104; VCPVB5.
CC HSSP; P30129; 4DPV.
CC InterPro; IPR001403; Parvo_coat.
CC Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 673 COAT PROTEIN VP2.
FT CHAIN 138 673 COAT PROTEIN VP3.
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 163 187 GLY-RICH.
SQ SEQUENCE 673 AA; 75103 MW; 5F244642B2214831 CRC64;

Query Match 17.1%; Score 556; DB 1; Length 673;
Best Local Similarity 25.0%; Pred. No. 2e-29;
Matches 160; Conservative 98; Mismatches 239; Indels 144; Gaps 19;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQPAKRLNFGQTGDSVSPDPQLGEP--PATP 58
Db 100 TSKGGDRALKRKLKYFARSNKGAKKANREPAPSTSNQQNMEVSDIPNDEAGNQPIELATR 159
QY 59 AAVGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTVNNH 118
Db 160 SVGSGSVGGG-----RGGSGVGYSTGGTGTTFSENIVVTNTRQFCIDKNGH 211
QY 119 LYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSFPRWQRLNNWGFPRKRLNFK 178

Db 212 LYKS-EVLNTGDTAHRQY-AITTPWSYFNENQYSSHSFSPNDWQHLVNDYERFRPKAMIVR 269
QY 179 LENIQKVEVTNDGVTTIANN-LTSTVQVFSDSYQLPYVLGSAHQGCLPPPPADVFMIP 237
Db 270 VYNLQIKQIMTDGAMGTVYNNDLTAGMHIFCDGHRYPYVQHPWDDQCMPELNSIWELP 329
QY 238 QYGYLTL-----NNGSQAVGRS-----SFYCLEYFPQMLRTGNNTFTSYTFEEVPFHS 286
Db 330 QYAYIPAPISVVDNNTTNTVEEHLKGVPLYMLENSDHEVLRNG----- 373
QY 287 SYAHSQSLDRMLNPLIDQYLYLNRTQN-----QSGSAQNKDLLFSRGSPP----- 331
Db 374 -----RIYRIYIQLWRLMRDRKQHHIQAASDDVQSTGQKQKNLLIQRTKQPNKQRF 424
QY 332 AGMSVQPKNWLPGPCYRQORVSKTKTDNNSNFTWTGASKYNLNGRESIINPGTAMASHK 391
Db 425 QNAALRTSNWMSGP-----GIARGTHNATLQTSAGALVTMTV 462
QY 392 DDEKDFPMSGVM-----IFGKE-----SAGASNTALDNVMTID 426
Db 463 NGAD---VSGVRAVRVGYSTDPIYGGQOPESDLLRLRYSASAEGQQNPILN----- 512
QY 427 EEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWIWAKIP 486
Db 513 -----AARHTFTREATKLITGNSGADGDYKEWMLPNQMWDAPISRYNPVWVKVP 564
QY 487 HTDGHFHPSPMLMGFGFLKHPPPQILIK--NTPVPANPPAEFSATKFAFITQYSTQVSV 544
Db 565 RVNRKTLTLDTQDSIPMSHPPGTIFIKLARIPVPGNGD-----SFLNIYVTGQVSC 615
QY 545 EIEWELQKENSKRWNPEVQYTSNYAKSANVD-FTVDNNGLY 584
Db 616 EVVWEVEKRGTKNWRPEYMHs---ATNMSVDAYTINNAGVY 653

RESULT 3
COAT_PAVHB STANDARD; PRT; 781 AA.
AC P07299;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Probable coat protein VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate AU;
RX MEDLINE=86200451; PubMed=3701931;
RA Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;
RT "Nucleotide sequence and genome organization of human parvovirus B19
isolated from the serum of a child during aplastic crisis.";
RL J. Virol. 58:921-936(1986).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M13178; AAA66867.1; -;
CC PIR; A24299; VCPV19.
CC InterPro; IPR001403; Parvo_coat.
CC Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 781 AA; 86015 MW; 8C6254DBD0576B07 CRC64;

Query Match 15.2%; Score 495; DB 1; Length 781;
Best Local Similarity 27.3%; Pred. No. 2.7e-25;
Matches 148; Conservative 84; Mismatches 257; Indels 54; Gaps 16;

QY 51 LGEPPATPAAGPTTMAAGGAPMADNNEGADGVGNASGNWCHDSTWLGLDRVITSTRTW 110
Db 213 LPEVPAYNASEKYPSTMTVNSAE-ASTGAGGGGNSVKSWMSEGAFTSANSVTCTFSRQF 271
QY 111 ALPTYNNHLYKQISSASTGASND-----NHFGYSTPWGYFDNFRFCHFSPRDW 160
Db 272 LIPYDPEHHYKVFSPAASSCHNASGKEAKVCTISPIMGYSTPWRYLDFNALNLFSPLEF 331
QY 161 QRLINNNGFRPKRLNFKLNIQKVT--TNDGVTTIANNLSTVQVFSDEYQLPYVL 218
Db 332 QHLIENYGSIAPDALTVTIIEIAVKDVTDKTGGV-QVTDSTTGRCLMLVDHEYKYPYVL 390
QY 219 GSAHQGLPPFPADVFMIPOGYLTLNN-GSQAVG-----RSSFYCLEYFPFSQMLR 268
Db 391 GQGQDTLAPELPIWVFPPOYAYLTVGDVNTQGISGDSKKLASEESAFYVLEHSSFQLLG 450
QY 269 TGNNTFSYTFEEVPHSSYAHSQSLDRLMNPIDQYLYLNRTQNQSGSAQNKDLLFSR 328
Db 451 TGGTASMSYKFPVPVPPENLEGCQHFYEMYNPL---YGSRLGVPTLGGDPKFRSL---- 503
QY 329 GSPAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPG-TAM 387
Db 504 -THEDHAIQPNFMPGPLVNSVSTKEGSSNTGAGKALTGLSTGTSQNTRIISLRPGPVSQ 562
QY 388 ASHKDDDEDKFFPMGSGVMIFGKESAGASNTALDNV-----MITDEEIKATNPVATERF 440
Db 563 PYHHWTDKYVTGINAISHGQTYG---NAEDKEYQGVGRFPNEKEQLKQLGLNMHTY 619
QY 441 GTVAVNFQSSSTDPAATGVHAMGALPGMWQDRDVLQGPINAKIPHTDGHFHS-PLMG 499
Db 620 -----FPNKGTOQYTDQIE-RPLMVGSVWNRRLHYESQLWSKIPNLDDSEFKTQFAALG 672
QY 500 GFGLKNPPPPQILIKNTVPVPAAPPAEFSATKFAFITQYSTGQVSVIEWEL-QKENSKRW 558
Db 673 GWGLHQPPPPQIFLK--ILPQSGPIGGIKSMGITTLLVQYAVGIMTVMTFKLGRKATGRW 730
QY 559 NPE 561
Db 731 NPQ 733

RESULT 4
COAL_PAVP9 STANDARD; PRT; 729 AA.
AC P33484;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Porcine parvovirus (strain 90HS) (ppv).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=33725;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89319168; PubMed=2750278;
RA Sakurai M., Nishimori T., Ushimi C., Nakajima H.;
RT "Nucleotide sequence of capsid protein gene of porcine parvovirus.";
RL Virus Res. 13:79-86(1989).
CC -!- SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED
CC VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.
CC -!- MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT
CC OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
DR PIR; A60006; A60006.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
```

```
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 729 COAT PROTEIN VP1.
FT CARBOHYD 151 729 COAT PROTEIN VP2.
FT CARBOHYD 172 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 198 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 282 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 330 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 433 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 471 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 573 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 604 729 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 651 729 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 729 AA; 80938 MW; 92538BFF9A0C78E6 CRC64;

Query Match 7.8%; Score 254.5; DB 1; Length 729;
Best Local Similarity 22.7%; Pred. No. 2e-09;
Matches 151; Conservative 95; Mismatches 277; Indels 141; Gaps 33;

QY 9 VEQSPQEPDSSSGIGK-----TGQOPAKKRLNFGQTGDSSEVPDQPLGEPPATP 58
Db 109 VRRSPRKHPGSKPPGKRPAPRHIFINLAKKAKAGTSNTNSMSSENVQHNPINAGTEL- 167
QY 59 AAVGPTTMAAGGAPMADNNEGADGVGNASGNWH--CDSTWLGD---RVITSTRTWALP 113
Db 168 SATGNESGGGGGGG---GGRGAGGVGVSTGSFNNQTEFOYLGEGLVRTAHASRLHLN 223
QY 114 TYNHLYKQIS--SASTGAS---NDNHFGYSTPWGYFDNFRFCHFSPRDWQRLINN 167
Db 224 MPEHETYKRIHVLNSESAGVAGQVQDDAHTQMTVPWSLIDANAWGVWFNPADWQLISNM 283
QY 168 WGFRRPKRLNFKLNIQKVT---TNDGVTTIANNLSTVQVFSDEYQLPYVLGSAHQ 224
Db 284 TEINLVSEFEQELFNVLKTTITESATSPPTKIYNNDLTASLMVALDTNNTLPYTPAAPRSE 343
QY 225 CLPPFPADVEMIPQYGYL-----TLNNGSQAVGRS-----SPCYCLE-YFPS 264
Db 344 TLGYFPLWPKPTQYRYLSCTRNLPPTYTGSQQTDSIQTGLHSDIMFYTIENAVPI 403
QY 265 QMLRTGNMF-TFSYTFEEVPHSSYAHSQSLDR-----LMNPLI--DQYLYLNRTQ 313
Db 404 HLLRTGDEFSTGIYHFDTKPL--KLTHSWQTNRSGLPPKLLTEPTTEGDQHPGTLPAAN 461
QY 314 NQSGSAQNKDLLFSRSGSPAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFTWTGASKYN 373
Db 462 TRKGYHQTWNNSYTEAT---AIRP-----AQVGYNTPYMNFYSNGGPP- 502
QY 374 LNGRESIINP--GTAMASHKDDKFFPMGSGVMIFGKESAGASNTALDNVMITDEEEIKA 431
Db 503 -----LTPIVPTADTQYNDDE----PNGAIRFTMGYQHGLTTS-----SQELERYT 545
QY 432 TNP-----VATERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQ 471
Db 546 FNPQSKCGRAPKQFQFNQQAPLNLENTNNGTLLPSDPIGGKPNMHFMTLNTYGLTALNN 605
QY 472 DRDYYLQGPINAKIPHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPVPAAPPAEFA- 527
Db 606 TAPVFPNGQIWDKELDTDLKPRLH---VTAPFVCKNNPPGQLFVKIAP---NLTDDEFNAD 659
QY 528 TKFASFITQYSTGQVSVIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEP 587
Db 660 SPQQPRIITYSNFWKGTLTFTAKMRSSNMWNPQQHTT-----TAENIGNYI-P 708
QY 588 RPIG 591
Db 709 TNIG 712

RESULT 5
COAL_PAVPN
ID COAL_PAVPN STANDARD; PRT; 729 AA.
AC P18546; P22964; Q89816;
DT 01-NOV-1990 (Rel. 16, Created)
```


DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Porcine parvovirus (strain NADL-2) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10797;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085785; PubMed=2596019;
RA Vasudevacharya J., Basak S., Srinivas R.V., Companys R.W.;
RT "Nucleotide sequence analysis of the capsid genes and the right-hand
RT terminal palindrome of porcine parvovirus, strain NADL-2.";
RL Virology 173:368-377(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021005; PubMed=2219713;
RA Vasudevacharya J., Basak S., Srinivas R.V., Companys R.W.;
RT "The complete nucleotide sequence of an infectious clone of porcine
RT parvovirus, strain NADL-2.";
RL Virology 178:611-616(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90010964; PubMed=2794971;
RA Ranz A.I., Manclus J.J., Diaz-Aroca E., Casal J.I.;
RT "Porcine parvovirus: DNA sequence and genome organization.";
RL J. Gen. Virol. 70:2541-2553(1989).
CC -1- SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED
CC VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.
CC -1- MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT
CC OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32787; AAA46917.1; -
DR EMBL; M32787; AAA46918.1; -
DR EMBL; M38367; AAA46919.1; -
DR EMBL; M38367; AAA46921.1; -
DR EMBL; D00623; BAA00502.1; -
DR PIR; B33302; VCPVPP.
DR PIR; B33743; VCPVNA.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 729 COAT PROTEIN VP1.
FT CHAIN 151 729 COAT PROTEIN VP2.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 175 190 GLY-RICH.
FT CONFLICT 10 10 G -> KGSGFGVVAYILQIFLYITG (IN REF. 3).
FT CONFLICT 16 18 YKY -> TI (IN REF. 3).
FT CONFLICT 56 56 Y -> T (IN REF. 3).
FT CONFLICT 164 164 G -> A (IN REF. 3).
FT CONFLICT 195 195 T -> S (IN REF. 3).
FT CONFLICT 242 242 V -> S (IN REF. 3).
FT CONFLICT 294 294 E -> A (IN REF. 3).
FT CONFLICT 705 705 N -> K (IN REF. 3).
FT CONFLICT 715 715 R -> K (IN REF. 3).
SQ SEQUENCE 729 AA; 80946 MW; EF816E246C80DB42 CRC64;

Query Match 7.8%; Score 254.5; DB 1; Length 729;
Best Local Similarity 22.7%; Pred. No. 2e-09;
Matches 152; Conservative 94; Mismatches 272; Indels 151; Gaps 34;
QY 9 VEQSPQEPDSSSGIGK-----TGQQPAKKRLNFGQTGSESVPDPQPLGEPPTP 58
DB 109 VRRSPRKHGSKPPGKRPAPRHFIFINLAKKAKGTSNTNSNSMSENVQHNPNAGTEL- 167
QY 59 AAVGPTTASGGGAPMADNNEGADGVGNASG--NWHCDSTWLGD---RVITTSTRTWALP 113
DB 168 SATGNESGGGGGGG---GGRGAGGVGVSTGTENNQTEFQYLGEGLVRIATAHASRLIHLN 223
QY 114 TYNNHLYKQIS--SASTGAS---NDNHYFGYSTPWGYFDNRFHCHFSRDMQRLINN 167
DB 224 MPEHETKYRIHVLNSESGVAGQMVQDDAHTQMVTWPSLIDANAWGVNFPADWQLISNNM 283
QY 168 WGFPRKRLNFKLFIQKVEVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQG 224
DB 284 TEINLVSFQEIEFNVVLKTTITESATSPPTKIYNNDLTASLMVALDTNNTLPYTPAAPRSE 343
QY 225 CLPPFPADVFMIPQYGYL-----TLNNGSQAVGRS-----SPYCLE-YFPS 264
DB 344 TLGFPYPLPTKPTQYRYLLSCIRNLNPPYTGQSQQITDSIQTGLHSDIMFYTIENAVPI 403
QY 265 QMLRTGNPF-TFSYTFEEVPHSSYAHSSQSLDR-----LMNPLI--DQYLYLNRTO 313
DB 404 HLLRTGDEFSTGIYHFDTKPL--KLTHSWQTNRSGLPLPKLLTEPTTEGDQHPGTLPAAN 461
QY 314 NQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGCYRQQRVSKTKTDNNSNFTWTGASKYN 373
DB 462 TRKGYHTQINNSYTEAT---AIRP-----AQVYNTPYMNFYSNGGPP- 502
QY 374 LNGRESIINP--GTAMASHKODEDKFFPMGVMIFGKESAGASNTALD---NVMITDEEE 428
DB 503 -----LTPIVPTADTQYNDDEPN-----GAIRFTMDYQHGHLTSSQE 540
QY 429 IK--ATNP-----VATERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP- 466
DB 541 LERYTFNPQSKCGRAPKQFNQQAPLNLENTNNGTLLPSDPIGKSNMHEFMTLNTYGPL 600
QY 467 GMVWQDRDVLQGIWAKIPHTD--GHFHPSPLMGGFGLK-NPPPQILIKNTVPVPANPPA 523
DB 601 TALNNTAPVFPNGQIWDKELDTDLKPRLH---VTAPFVCKNPPGQLFVKIAP---NLTD 654
QY 524 EFSA-TKFASFITQYSTQVSVEIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNNG 582
DB 655 DFNADSPQOPRIITYSNFWWKGLTFTTAKMRSSNMWNPIQOHTT-----TAENIG 704
QY 583 LYTEPRPIG 591
DB 705 NYI-PTNIG 712

RESULT 6

COAL_PAVPK STANDARD; PRT; 729 AA.
ID COAL_PAVPK AC P52501;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Porcine parvovirus (strain Kresse) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=73487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96183900; PubMed=8642680;
RA Bergeron J., Hebert B., Tijssen P.;
RT "Genome organization of the Kresse strain of porcine parvovirus:
RT identification of the allotropic determinant and comparison with
RT those of NADL-2 and field isolates.";
RL J. Virol. 70:2508-2515(1996).

CC -|- SUBUNIT: MATURE VIRION CONTAINS THREE CAPSID PROTEINS DESIGNATED
CC VP1, VP2, AND VP3 AND A NONCAPSID PROTEIN NS-1.
CC -|- MISCELLANEOUS: VP3 MIGHT BE A POSTTRANSLATIONAL CLEAVAGE PRODUCT
CC OF VP2 IN SEVERAL AUTONOMOUS PARVOVIRUSES.
CC -|- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U44978; AAC40230.1; --
CC EMBL; U44978; AAC40231.1; --
CC HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 729 COAT PROTEIN VP1.
FT CHAIN 151 729 COAT PROTEIN VP2.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 729 AA; 80835 MW; B6345BFA0568A1F6 CRC64;

Query Match 7.8%; Score 252.5; DB 1; Length 729;
Best Local Similarity 22.7%; Pred. No. 2.7e-09;
Matches 151; Conservative 95; Mismatches 277; Indels 141; Gaps 33;

Qy 9 VEQSPQEPDSSSGIGK-----TGQPAKRLNFGQTGDSSEVPDPQPLGEPPTATP 58
Db VRRSPRKHPGSKPPGKRPAPRPHIFINLAKKAKGTSNTNSMSSENVEQHNPINAGTEL- 167

Qy 59 AAVGPTTMASSGGGAPMADNNEGADGVGNASGNWH--CDSTWLGD--RVITTSTRTWALP 113
Db SATGNESGGGGGGG---GGRGAGGVGVSTGTFNNQTEFQYLGELVRITAHASRLIHLN 223

Qy 114 TYNHLYKQIS--SASTGAS---NDNHFGYSTPWGYFDFNRPFHCHFSPRDWQRLNNN 167
Db MPEHETTKRIHVLNSESAGVAGQVQDDAHTQMTWPWSLIDANAGVWFENPADWQLISNM 283

Qy 168 WGFPRKRLNFKLNIQKVEV---TNDGVTTIANLSTVQFSDSEYQLPYVLGSAHQG 224
Db TEINLVSEFEQEIFNVVLKTTITESATSPPTKIYNDLTASLMVALDTNNTLPYTPAAPRSE 343

Qy 225 CLPPFPADVFMIPQYGL-----TLNNGSQAVGRS-----SFYCLE-YFPS 264
Db TLGFYPLPTKPTQYRYLLSCTRNLNPPPTYTGQSQQITDSIQTLHSDIMFYTIENAVPI 403

Qy 265 QMLRTGNF-TFSYTFEEVPPFHSSYAHQSQSLDR-----LMNPLI--DQYLYLNRQTQ 313
Db HLLRTGDEFSTGIYHFDTKPL--KLTHSWQTNRSGLPPLKLTTEPTTEGDQHPGTLPAAN 461

Qy 314 NQSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYN 373
Db TRKGHYQTINNSYTEAT---AIRP-----AQGVNTPYMNFEYSNGGPF- 502

Qy 374 LNGRESINP--GTAMASHKDDKDFPMSGVMI FGKESAGASNTALDNVMITDEEEIKA 431
Db -----LTPIVPTADTQYNDDE----PNGAIRFTMGYQHGLTTS-----SQELERYT 545

Qy 432 TNP-----VATERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQ 471
Db FNPQSKCGRAPKQFNQQAPLNLENTNNGTLLPSDPIGGKENMHFMTLNTYGPALTALNN 605

Qy 472 DRDVYLQGPWAKIPHTD--GHFHPSPMLGGFGLK-NPPPPQILIKNTVPVPANPPAEPFA- 527
Db TAPVFPNGQIWDKELDTLKPRLH---VTAPFVCKNNPPGQLFVKIAP---NLTDDEFNAD 659

Qy 528 TKPASFITQYSTGQSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEP 587
Db SPQPRIITYSNFWWKGTLTFTAKMRSSNMWNPIQQHTT-----TAENIGNYI-P 708

Qy 588 RPIG 591
Db 709 TNIG 712

RESULT 7
COAT_PAVHH STANDARD; PRT; 722 AA.
AC P03136;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
genes by hybrid-arrested translation."
RL J. Virol. 45:173-184(1983).
CC -|- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01457; CAB57285.1; ALT_SEQ.
DR PIR; A03699; VCPVV2.
DR HSSP; P07302; 1MVM.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 722 COAT PROTEIN VP1.
FT CHAIN 131 722 COAT PROTEIN VP2.
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 647 647 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 675 675 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 155 170 GLY-RICH.
SQ SEQUENCE 722 AA; 79737 MW; 6BB678391AA5DC31 CRC64;

Query Match 7.5%; Score 244; DB 1; Length 722;
Best Local Similarity 21.4%; Pred. No. 9.7e-09;
Matches 142; Conservative 94; Mismatches 267; Indels 160; Gaps 29;

Qy 8 PVEQSPQEPDSSSGIGKGTQQ---PA-----KKRLNFGQTGDSSEVPDPQPLGEP 54
Db PKLSTDSEP-GTSGVSRPPGKRTKPPAHIFVNQRAKKKRAASLAAQQRTLTMSDGTETNPQ 140

Qy 55 PATPAAVGPTTMASSGGGAPMADNNEGADGVGNASGNWHCDSTW--LGDRVITTTFTWAL 112
Db DTGIANARVERSADGGGS-SGGGGSGGGGIGVSTGYDNTQTTYKFLGDGWVEITAHASRL 199

Qy 113 -----PTYNHLYKQISSASTGASNDNHFGYSTPWGYFDFNRFHCHFSPR 158
Db LHLGMPPSENYCRVTVHNNQTTGHGTVKVGKNMAYDTHQQIW--TPWSLV DANAWGVWFQPS 258


```
FT STRAND 652 668
FT STRAND 683 683
FT STRAND 691 691
FT HELIX 692 694
SQ SEQUENCE 718 AA; 79017 MW; B43C1762ED5F74B9 CRC64;

Query Match 6.6%; Score 214.5; DB 1; Length 718;
Best Local Similarity 21.0%; Pred. No. 8.6e-07;
Matches 142; Conservative 89; Mismatches 273; Indels 171; Gaps 30;

Qy 8 PVEQSPQEPDSSSGIGKTG-----QPAKKRLNFG---QTGDSSEVPDPQLG 52
Db 83 PKLATDSEP-GTSGVSRAGKRTPPAYIFINQARAKKLTSSAAQSSQTMDSGTSQPDG 141
Qy 53 -----EPPATPAAGVPTTMASGGGAPMADNNEGADGVGNASGNWHCDS--TWLGD--R 101
Db 142 GNAVHSAARVERAADGPG--GSGGG-----GSGGGGVGVSTGSDYNQTHYRFLGDGWE 193
Qy 102 VITSTRTRTVALPTYNHLYKQI-----SSASTGASNDNHVFGYSTPWGYDFENRFCH 154
Db 194 ITALATRLVHLNMPKSENYCRIRVHTTDTSVKGNMAKDDAHEQIWPWSLVDANAWGVW 253
Qy 155 PSPRDQWRQLNNWGFPRKRLNFKLFNIQVKEVTND----GVTTIANNTLSTVQVPSDS 210
Db 254 LQPSDWQYICNTMSQLNLVSLDQEIFNVVLKTVTEQDSGGQAIKIYNNDLTACMMVAVDS 313
Qy 211 EYQLPVVLGSAHQGLPPFPADVFMIPOGY-----LTLNN-----GSQAV 251
Db 314 NNILPYTPAANSMETLGFYPWKPTIASPYRYFCVDRDLSVTYENQEGTIEHNVMGTFKG 373
Qy 252 GRSSFCYLEYFPPS-QMLRTGNPF-TFSYTFEEVFFHSSVAHSQSLDRLMN--PLIDQYLY 307
Db 374 MNSQPTIENTQQITLLRTGDEFATGYFDTNPV--KLTHWTQTNRLQGPPLLTSTF-- 429
Qy 308 YLNRTQNSGSAQNKOLLFSRGSAGMSVQPKNL-----P 343
Db 430 --PEADTDAGT-----LTAQGSRHGATQMEVNVWVSEAIRTPAQVGFQCPHNDFEASRA 481
Qy 344 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDF----- 397
Db 482 GP-FAAPKVPADVTQGVDRANGSVRYSGKQGHENWAAHGPAPERYTWTDETNFGSGRDT 540
Qy 398 ---FPMGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQS---SST 452
Db 541 RDGFIQSAPLVVPPPLNGILTNA-----NPIGTKN-----DIHFSNVFNSY 581
Qy 453 DPATGDVHAMGALPGMVWQDRDVLQGPWAK---IPHTDGHFSPMLMGFGFLGNPPPPQ 509
Db 582 GPLTAFSH-----PSPVYPQGIWDKELDLHKPRLHITAPFV---CKNNAFGQ 627
Qy 510 ILIKNTP--VPANPPAEFSATKPFASFIQYSTGQSVSEIWELOKENSKRWNPEVQYTSN 567
Db 628 MLVRLGPNLTDQYDPNGATLSRIVTYGTFFWKKGKLTWRAKLRA-----NTTNVPVYQ--- 679
Qy 568 YAKSANVDFTVDNNG 582
Db 680 -----VSVEDNG 686
```

```
RESULT 12
COAT_MEVA
ID COAT_MEVA STANDARD; PRT; 722 AA.
AC P27437;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Mink enteritis virus (strain Abashiri) (MEV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91202123; PubMed=2016597;
```

```
RA Kariatsumari T., Horiuchi M., Hama E., Yaguchi K., Ishiguro N.,
RA Goto H., Shinagawa M.;
RT "Construction and nucleotide sequence analysis of an infectious DNA
RT clone of the autonomous parvovirus, mink enteritis virus.";
RL J. Gen. Virol. 72:867-875(1991).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00765; BAA00663.1; -.
DR PIR; B38350; VCPVME.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 722 COAT PROTEIN VP1.
FT CHAIN 139 722 COAT PROTEIN VP2.
FT DOMAIN 160 177 GLY-RICH.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 581 581 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 643 643 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 722 AA; 79823 MW; 9DADBCAB2EF9F622 CRC64;

Query Match 6.6%; Score 214; DB 1; Length 722;
Best Local Similarity 20.2%; Pred. No. 9.3e-07;
Matches 139; Conservative 101; Mismatches 259; Indels 188; Gaps 34;

Qy 7 RPVEQSPQEPDSSSGIGKTGQPAKKRLNFGQTGDSSEVP---DPQLGEBPPA--TPAA 60
Db 105 KPTKRSKPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAQVQPDGGQPAVRNERA 158
Qy 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWH-----CDSTWLGDRVITSTRTWALP 113
Db 159 TSGSGNGSGGG-----GGSGGVGISTGFNNQTEFKLENGWV--EITANSSRLVHLN 210
Qy 114 TYNHLYKQI-----SSASTGASNDNHVFGYSTPWGYDFNRFCHESPRDQWRQLNN 166
Db 211 MPESENYKRVVNNMDKTAVKGNMALDDTHVQIVTPWSLVDANAWGVWFNPGDWQLIVNT 270
Qy 167 NWGFRPKRLNFKLFNIQVKEVT---TNDGVTTIANNTLSTVQVPSDSEYQLPYVLGSAHQ 223
Db 271 MSELHLVSPQEIEFNVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPTPAAMRS 330
Qy 224 GCL-----PPPPA-----DVFTIPOYGYLTNLNGSQAVGRS-----SFYCL 259
Db 331 ETLGYFPWKPTIPTPWRYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDDVQFYTI 384
Qy 260 E-YFPSSQMLRTGNMF-TFSYTFEEVFPFHSSVAHSQSLDRLMN--PLIDQYLYLNRTONQ 315
Db 385 ENSVPVHLRLTGDEFATGTFEFDCKP--CRLTHWTQTNRALGLPP-----FLNSLPQS 435
Qy 316 SGSAQNKDLFSRGSAGMS-----VQPKN-WLPGPCYRQQRVSK----- 354
Db 436 EGATNFGDIGVQDQKRRGVQTMGNTDYITEATIMRPAEVGYSAPYSPPEASTQGPFTPI 495
Qy 355 -----TKTDNNSNFTWTGASKY---NLNGRESIINPGT---AMASHKDDKDEKFFPMS 401
Db 496 AAGRGAQTDENQAA---DGDPRYAFGRGHQKQKTTTGETPERFTYIAHQT----- 544
Qy 402 GVMIFGKESAGASNTALD-NVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVH 460
Db 545 -----GRYPAGDWIQNINFNLPVTNDNVLLPTDPIG----GKTGINY--TNIFNTYGPLT 593
```


-!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M38245; AAB02799.1; -
EMBL; M38245; AAB02800.1; -
HSSP; P30129; 4DPV.
InterPro; IPR001403; Parvo_coat.
Pfam; PF00740; Parvo_coat; 1.
Coat protein; Glycoprotein.
CHAIN 1 727 COAT PROTEIN VP1.
CHAIN 144 727 COAT PROTEIN VP2.
CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 207 207 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 660 660 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 727 AA; 80342 MW; 115F3E1A79098EBE CRC64;

Query Match 6.4%; Score 208; DB 1; Length 727;
Best Local Similarity 19.8%; Pred. No. 2.3e-06;
Matches 137; Conservative 96; Mismatches 260; Indels 200; Gaps 31;

QY 7 RPVEQSPQEPDSSSGIGKTGQPAKKRLNFGQTGDSSEVP---DPQLGEPPEA--TPAA 60
Db 110 KPTKRSPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAQPDPGGQPAVRNERA 163
QY 61 VGPTTMASSGGGAPMADNNEGADGVGNASGNWH-----CDSTWLGDRVITTTSTRTWALP 113
Db 164 TSGSGNGSGGG-----GGSGGVGISTGTFTNNQTEFKFLENGWV--EITANSSRLVHLN 215
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLINN 166
Db 216 MPESENYRRVVNNMDKTAVNGNMALDDIHAQIVTPWSLVDANAWGVWFPNGDWQLIVNT 275
QY 167 NWGFRPKRLNFKLFNIQKVEV---TNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 223
Db 276 MSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPTPAVRS 335
QY 224 GCL-----PPFPA-----DVFMIPOGYLTLNNGSQAVERS-----SFYCL 259
Db 336 ETLGFYPWKPTIPTPWRYFFQWDRTLIPSH-----TGSGTPTNIYHGTDPPDDVQFYTI 389
QY 260 E-YFPSQMLRTGNPF-TFSYTFEEVPFHSSYAHQSQSLRLMN--PLIDQYLYLNRTQNG 315
Db 390 ENSVPVHLRLTGDEFATGTFFDCKP--CRLTHTWQTNRALGLPP-----FLNSLPQS 440
QY 316 SGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF-----TW 366
Db 441 EGATNFGDI-----GV-----QQDKRGRVGTQMGNTNYITEATIMRPAE 478
QY 367 TGASKYNLNGRESIINP-----GTAMASHKDDKEDKFFPMGSMVIFGKESAGASNTALD-- 419
Db 479 VGYSAPYYSFEASTQGPFTPIAAGRGAQTDENQAADGNRYAFGRQHGQKTTTGTETP 538
QY 420 -----NVMITDEEIKATNPVATERFGTVAVNFQSSSTDP 454
Db 539 ERFTYIAHQDTGRYPEGDWIONINFLPVTNDNVLLPTDPDG-----GKTGINY--TNIFN 592
QY 455 ATGDVHANGALPGMVWQDRDVYLGQPIWAKIPHTDGHFHSPLMGGFGLK----- 504
Db 593 TYGPLTALNNVP-----PVYPNGQIWDKEFDTD-----LKPRLHVNAFPV 632
QY 505 ---NPPPIILIKNTPVPAN---PPAEFSATKFASFITQYSTGQVSVEIEWELQENSKRW 558

Db 633 CQNNCPGQLFVKVAPNLITNEYDPDASANMSR-----IVTYSDFWMWKGKLVFKAKLRASHTW 688
QY 559 NPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 689 NPIQQMSIN-----VDNQFNYY-PSNIG 710

RESULT 15
COAT_PAVCD STANDARD; PRT; 737 AA.
AC P17455;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Canine parvovirus (strain CPV-D Cornell 320) (CPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10790;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89020796; PubMed=3176341;
RA Parrish C.R., Aquadro C.F., Carmichael L.E.;
RT "Canine host range and a specific epitope map along with variant sequences in the capsid protein gene of canine parvovirus and related feline, mink, and raccoon parvoviruses.";
RL Virology 166:293-307(1988).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M23255; AAA47158.1; -
DR PIR; A31163; VCPVCD.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 737 COAT PROTEIN VP1.
FT CHAIN 154 737 COAT PROTEIN VP2.
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 658 658 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN 175 192 GLY-RICH.
SQ SEQUENCE 737 AA; 81504 MW; 39FE81D3E5435EC9 CRC64;

Query Match 6.4%; Score 208; DB 1; Length 737;
Best Local Similarity 19.8%; Pred. No. 2.4e-06;
Matches 137; Conservative 96; Mismatches 260; Indels 200; Gaps 31;

QY 7 RPVEQSPQEPDSSSGIGKTGQPAKKRLNFGQTGDSSEVP---DPQLGEPPEA--TPAA 60
Db 120 KPTKRSPPPHIFINLAK-----KKKAGAGQVKRDNLAPMSDGAQPDPGGQPAVRNERA 173
QY 61 VGPTTMASSGGGAPMADNNEGADGVGNASGNWH-----CDSTWLGDRVITTTSTRTWALP 113
Db 174 TSGSGNGSGGG-----GGSGGVGISTGTFTNNQTEFKFLENGWV--EITANSSRLVHLN 225
QY 114 TYNHLYKQI-----SSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDWQRLINN 166
Db 226 MPESENYRRVVNNMDKTAVNGNMALDDIHAQIVTPWSLVDANAWGVWFPNGDWQLIVNT 285
QY 167 NWGFRPKRLNFKLFNIQKVEV---TNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQ 223

Db 286 MSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPTPAAMRS 345
QY 224 GCL-----PPFPA-----DVFMIPQYGYLTLNNGSQAVGRS-----SFYCL 259
Db 346 ETLGFPYWKPTIPTPWRYYPQWDRTLIPSH-----TGTSGTPTNIYHGTDPPDDVQPYTI 399
QY 260 E-YFPSQMLRTGNF-TFSYTFEEVPPFHSSYAHSQSLDRLMN--PLIDQYLYLNRTQNO 315
Db 400 ENSVPVHLLRTGDEFATGTFEFDCKP--CRLTHWTQNRALGLPP-----FLNSLPQS 450
QY 316 SGSAQNKDLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNSNF-----TW 366
Db 451 EGATNFGDI-----GV-----QQDKRGGVTQMGNTNYITEATIMRPAE 488
QY 367 TGASKYNLNGRESIINP-----GTAMASHKODEDKFFPMMSGVMIFGKESAGASNTALD-- 419
Db 489 VGYSAPYYSFEASTQGPFKTPIAAGRGAQTDENQAADGNPRYAFGRQHGQKTTTGETP 548
QY 420 -----NVMITDEEEIKATNPVATERFGTVAVNFQSSSDP 454
Db 549 ERFTYIAHQDTGRYPEGDWIQNINFNLPVTNDNVLLPTDPIG---GKTGINY--TNIFN 602
QY 455 ATGDVHAMGALPGMVWQDRDVYLGQPIWAKIPHDTDGHFHPSPLMGGFGLK----- 504
Db 603 TYGPLTALNNVP-----PVYPNGQIWDKEFDTD-----LKPRLHVNPAPV 642
QY 505 ---NPPPPQILIKNTPVPAN---PPAEFSATKFASFITQYSTQGVSVIEWELQKENSQRW 558
Db 643 QONNCPGQLFVKVAPNLTNEYDPPDASANMSR-----IVTYSDFWWKGLVFKAKLRASHTW 698
QY 559 NPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 699 NPIQQMSIN-----VDNQFNIV-PSNIG 720

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:00:06 ; Search time 31.0877 Seconds
(without alignments)
4972.168 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL 23:*
1: sp_arChaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3251	100.0	736	12 Q9WBP8	Q9wbp8 adeno-assoc
2	3229	99.3	736	12 O56137	O56137 adeno-assoc
3	2832	87.1	736	12 O56139	O56139 adeno-assoc
4	2815	86.6	736	12 Q65311	Q65311 adeno-assoc
5	2764.5	85.0	735	12 O56652	O56652 adeno-assoc
6	2759.5	84.9	598	12 O56653	O56653 adeno-assoc
7	2709.5	83.3	737	12 Q8JQG0	Q8jqg0 adeno-assoc
8	2679	82.4	738	12 Q8JQF8	Q8jqf8 adeno-assoc
9	2481.5	76.3	533	12 Q92917	Q92917 adeno-assoc
10	1830.5	56.3	734	12 O41855	O41855 adeno-assoc
11	1717	52.8	732	12 Q67666	Q67666 goose parvo
12	1714	52.7	732	12 Q83290	Q83290 muscovy duc
13	1712	52.7	587	12 Q67667	Q67667 goose parvo
14	1712	52.7	732	12 Q8V395	Q8v395 goose parvo
15	1702	52.4	732	12 Q65444	Q65444 barbarie du
16	1701	52.3	732	12 Q83289	Q83289 muscovy duc

17	1700.5	52.3	724	12 Q9YIJ1	Q9Yij1 adeno-assoc
18	1698	52.2	534	12 Q67668	Q67668 goose parvo
19	1695	52.1	587	12 Q65445	Q65445 barbarie du
20	1676	51.6	534	12 Q65446	Q65446 barbarie du
21	1540	47.4	676	12 Q67672	Q67672 goose parvo
22	715	22.0	179	12 Q9W8U7	Q9w8u7 duck parvov
23	715	22.0	179	12 Q9WA24	Q9wa24 goose parvo
24	711	21.9	179	12 Q9WN18	Q9wn18 duck parvov
25	709	21.8	179	12 Q9WN19	Q9wn19 duck parvov
26	707	21.7	179	12 Q9WN20	Q9wn20 goose parvo
27	632.5	19.5	703	12 Q8QQV5	Q8qqv5 minute viru
28	618	19.0	571	12 Q8QQV4	Q8ggv4 minute viru
29	545.5	16.8	947	12 Q918U9	Q918u9 bovine parv
30	508	15.6	781	12 Q9PZT0	Q9pzt0 human parvo
31	503.5	15.5	781	12 Q8JN56	Q8jns6 human parvo
32	501	15.4	554	12 Q9PZS9	Q9pzs9 human parvo
33	501	15.4	781	12 Q9JGP8	Q9jgp8 human parvo
34	500.5	15.4	773	12 Q913X1	Q913x1 human parvo
35	500.5	15.4	781	12 Q8JYE3	Q8jye3 erythroviru
36	499.5	15.4	781	12 Q8JYD9	Q8jyd9 erythroviru
37	499.5	15.4	785	12 Q9J0X4	Q9j0x4 pig-tailed
38	499	15.3	781	12 P89318	P89318 human parvo
39	499	15.3	781	12 P89319	P89319 human parvo
40	498.5	15.3	781	12 P89317	P89317 human parvo
41	497.5	15.3	781	12 Q912B8	Q912b8 human eryth
42	497	15.3	769	12 Q9PZT4	Q9pzt4 human parvo
43	497	15.3	781	12 P90223	P90223 human parvo
44	497	15.3	781	12 Q85191	Q85191 human parvo
45	497	15.3	781	12 P90221	P90221 human parvo

ALIGNMENTS

RESULT 1

Q9WBP8 ID Q9WBP8 PRELIMINARY; PRT; 736 AA.
AC Q9WBP8;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=85106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99214338; PubMed=10196295;
RA Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;
RT "Gene therapy vectors based on adeno-associated virus type 1.";
RL J. Virol. 73:3994-4003(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Xiao W., Wilson J.M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF063497; AAD27757.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81375 MW; CFABFB9BD5CD0595 CRC64;

Query Match 100.0%; Score 3251; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 3.9e-227;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDPQPLGEPATPAA	60
Db	138	TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDPQPLGEPATPAA	197
QY	61	VGPTTASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTWALPTVNNHLY	120
Db	198	VGPTTASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTWALPTVNNHLY	257
QY	121	KQISSASTGASNDNHYFGYSTPWGYFDNFRFHCHFSRQWLNNNNGWFRPKRLNFKLF	180

Db 258 KQISSASTGASNDNHYFGYSTPWGYFDNRFHCHFSRPRDQWRLNNWGFPRKRLNFKLF 317
QY 181 NIQVKEVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPFPADVFMIPOYG 240
Db 318 NIQVKEVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPFPADVFMIPOYG 377
QY 241 YLTLNNGSAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMNP 300
Db 378 YLTLNNGSAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMNP 437
QY 301 LIDQYLYLNRNTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 438 LIDQYLYLNRNTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN 420
Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN 557
QY 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGP 480
Db 558 VMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGP 617
QY 481 IWAKIPHDTGHHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFPASFITQYSTG 540
Db 618 IWAKIPHDTGHHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFPASFITQYSTG 677
QY 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 678 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736

RESULT 2

O56137 PRELIMINARY; PRT; 736 AA.
AC O56137;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028704; AAB95450.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 99.3%; Score 3229; DB 12; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.5e-225;
Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSVSPDPQPLGEPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSVSPDPQPLGEPATPAA 197
QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGRVITTTSTRTWALPTYNNHLY 120
Db 198 VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGRVITTTSTRTWALPTYNNHLY 257
QY 121 KQISSASTGASNDNHYFGYSTPWGYFDNRFHCHFSRPRDQWRLNNWGFPRKRLNFKLF 180

Db 258 KQISSASTGASNDNHYFGYSTPWGYFDNRFHCHFSRPRDQWRLNNWGFPRKRLNFKLF 317
QY 181 NIQVKEVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPFPADVFMIPOYG 240
Db 318 NIQVKEVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPFPADVFMIPOYG 377
QY 241 YLTLNNGSAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMNP 300
Db 378 YLTLNNGSAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEDVPHSSYAHQSGLDRLMNP 437
QY 301 LIDQYLYLNRNTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 438 LIDQYLYLNRNTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 497
QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN 420
Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDKFFPMSGVMIFGKESAGASNTALDN 557
QY 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGP 480
Db 558 VMITDEEEIKATNPVATERFGTVAVNLQSSSTDPAATGDVHMALPGMVWQDRDVYLQGP 617
QY 481 IWAKIPHDTGHHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFPASFITQYSTG 540
Db 618 IWAKIPHDTGHHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFPASFITQYSTG 677
QY 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 678 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736

RESULT 3

O56139 PRELIMINARY; PRT; 736 AA.
AC O56139;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028705; AAB95452.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81906 MW; DD523331AD5F0D70F CRC64;

Query Match 87.1%; Score 2832; DB 12; Length 736;
Best Local Similarity 85.7%; Pred. No. 9.1e-197;
Matches 514; Conservative 34; Mismatches 50; Indels 2; Gaps 2;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSVSPDPQPLGEPATPAA 60
Db 138 TAPGKKRPVDQSPQEPDSSSGVGKSGKQPAKKRLNFGQTGDSVSPDPQPLGEPAPATS 197
QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGRVITTTSTRTWALPTYNNHLY 120
Db 198 LGSNTMASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGRVITTTSTRTWALPTYNNHLY 257
QY 121 KQISSASTGASNDNHYFGYSTPWGYFDNRFHCHFSRPRDQWRLNNWGFPRKRLNFKLF 180
Db 258 KQISSQS-GASNDNHYFGYSTPWGYFDNRFHCHFSRPRDQWRLNNWGFPRPKLSFKLF 316

QY	181	NIQVKEVTNDGVTTIANNLSTVQVFS	SEYQLPYVLGSAHQCLPPFPADVFMIPQYG	241
Db	317	NIQVKEVTNDGVTTIANNLSTVQVFT	SEYQLPYVLGSAHQCLPPFPADVFMVPOYG	376
QY	241	YLTLNNGSOAVGRSSFCLEYFPSQMLRTGNN	FTFSYTFEEVPPFHSSVAHSQSLDRLMNP	300
Db	377	YLTLNNGSOAVGRSSFCLEYFPSQMLRTGNN	FQFSYTFEDVPFHSSVAHSQSLDRLMNP	436
QY	301	LIDQYLYLNRTO-NQSGSAQNKDLLFSRGS	PAGMSVQPKNWLPGPCYRQQRVSKTKTDN	359
Db	437	LIDQYLYLNRTOGTTSGTTNQSRLLFSQAGP	QSMQLQARNWLPQPCYRQQRLSKTANDN	496
QY	360	NNSNFTWTGASKYNLNGRESIINPGTAMASH	KDDEKFFPMGVMIFGKESAGASNTALD	419
Db	497	NNSNFPWTAASKYHLNGRDSLVNPGPAMASH	KDDEKFFPMHGNLIFGKEGTTASNAELD	556
QY	420	NVMITDEEEIKATNPVATERFGTVAVNFQSS	STDPAATGDVHAMGALPGMWQDRDVLQ	479
Db	557	NVMITDEEEIRTTNPVATEQYGTVANNLQSS	NTAPTTRTVNDQALPGMWQDRDVLQ	616
QY	480	PIWAKIPHTDGHFHPSPLMGGFGLKNPPPP	QILIKNTVPVPPANPPAEFSATKPFASFITQYST	539
Db	617	PIWAKIPHTDGHFHPSPLMGGFGLKHPPPP	QIMIKNTVPVPPANPTTFSAPKPFASFITQYST	676
QY	540	GQVSVEIEWELQKENS	KRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL	599
Db	677	GQVSVEIEWELQKENS	KRWNPVQYTSNYKSNVNDFTVDTNGVYSEPRPIGTRYLTRNL	736

RESULT 4

Q65311

PRELIMINARY;

PRT;

736 AA.

AC

Q65311;

DT

01-NOV-1996 (TrEMBLrel. 01, Created)

DT

01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT

01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE

Capaid protein.

OS

Adeno-associated virus 3.

OC

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

OX

NCBI_TaxID=46350;

RP

[1]

RP

SEQUENCE FROM N.A.

RC

STRAIN=3H;

RX

MEDLINE=96266430; PubMed=8661429;

RA

Muramatsu S., Mizukami H., Young N.S., Brown K.E.;

RT

"Nucleotide sequencing and generation of an infectious clone of adeno-associated virus 3.";

RL

Virology 221:208-217(1996).

RN

[2]

RP

SEQUENCE FROM N.A.

RC

STRAIN=3H;

RA

Muramatsu S., Brown K.E.;

RL

Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

DR

EMBL; U48704; AAC55049.1; -.

DR

InterPro; IPR001403; Parvo_coat.

DR

Pfam; PF00740; Parvo_coat; 1.

SQ

SEQUENCE 736 AA; 81660 MW; AFF1EP47B5C67A10 CRC64;

```

Db      258 KQISSQS-GASNDNHYFGYSTPWGYFDENRFHCHFSRDPWQRLLNNNWGFRPKCLSFKLF 316
Qy     181 NIQVKEVTNDGVTTIANNLSTTVQFSDSEYQLPYVLGSAAHQGLPPFPADVFMIPQYG 240
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     317 NIQVRGVTONDGTITIANNLSTTVQFTDSEYQLPYVLGSAAHQGLPPFPADVFMVPQYG 376
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     241 YLTLNNGSQAVGRSSFYCLEYPFSQMLRTGNNTFSPYTPEEVPHSSHSAHSQSLDLRLMNP 300
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     377 YLTLNNGSQAVGRSSFYCLEYPFSQMLRTGNNTFSPYTPEEDVFPHSSHSAHSQSLDLRLMNP 436
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     301 LIDQYLILNRTQ-NQSGSAQNKDLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDN 359
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     437 LIDQYLILNRTQGTTTSGTTNQSRLLFSQAQPQSMQLOARNWLPGPCYRQQRLSKTANDN 496
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     360 NNSNFWTGASKYNLNGRESIINPQTAMASHKDDEKFEPMSGVMIFGKESAGASNALD 419
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     497 NNSNFPTAASKYHLNGRDSLVPNPGPAMASHKDDEKFEPMHGNLI FGKEGTTASNAELD 556
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     420 NVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVYLOG 479
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     557 NVMITDEEIRTNPVATEQYGTVANLQSSNTAPTGTGVNHQCALPGMVWQDRDVYLOG 616
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     480 PIWAKIPHDTGHFHPSPLMGGFGLKNPPPQILIKNTPVPANPPAEFSATKFAFITOYST 539
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     617 PIWAKIPHDTGHFHPSPLMGGFGLKHPPPIQMINKNTPVPANPPTTFSPAKFAPITOYST 676
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Qy     540 QOVSVIEIEWELQENSKRWNPEVOYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||
Db     677 QOVSVIEIEWELQENSKRWNPEIQYTSNYNKSNVDFTVDTNMGVSEPRPIGTRYLTRNL 736
       |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||

RESULT 5
O56652 PRELIMINARY; PRT; 735 AA.
AC O56652;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major coat protein VP1.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
RT proteins affect viral infectivity: lack of an RGD integrin-binding
RT motif.";
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03780.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
FT VARIANT 76 76 D -> V.
FT VARIANT 553 553 D -> N.
FT VARIANT 567 567 T -> N.
FT VARIANT 677 678 HV.
FT VARIANT 710 710 V -> R.
SQ SEQUENCE 735 AA; 81944 MW; 980BEEF46908390B CRC64;
```

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSESVDPDPLGEPPTATPA

Db 138 TAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQPPAAPSG 197

QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLY 120

Db 198 LGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLY 257

QY 121 KOISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDQWRLNINNNGFRPKRLNFKLF 180

Db 258 KOISSQS-GASNDNHFGYSTPWGYDFDNRFCHFSFPRDQWRLNINNNGFRPKRLNFKLF 316

QY 181 NIQKVEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPQYG 240

Db 317 NIQKVEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYG 376

QY 241 YLTLNNGSOAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 300

Db 377 YLTLNNGSOAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSQSLDRLMNP 436

QY 301 LIDQYLYLNRNTQNSQSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360

Db 437 LIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496

QY 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMGSGVMIFGKESAGASNTALDN 420

Db 497 NSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLIFGKQSGSEKTNVDIEK 556

QY 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDYLQGP 480

Db 557 VMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATAVDNTQGVLPGMVWQDRDYLQGP 616

QY 481 IWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTG 540

Db 617 IWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTG 676

QY 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599

Db 677 QVSVEIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 6

O56653 PRELIMINARY; PRT; 598 AA.

ID O56653

AC O56653

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Major coat protein VP2.

OS Adeno-associated virus 2 (AAV2).

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

OX NCBI_TaxID=10804;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95088582; PubMed=7996133;

RA Ruffing M., Held H., Kleinschmidt J.A.;

RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid proteins affect viral infectivity: lack of an RGD integrin-binding motif.";

RL J. Gen. Virol. 75:0-0(0).

RN [2]

RP SEQUENCE FROM N.A.

RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,

RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,

RA Tratschin J.-D., Weitz M.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF043303; AAC03778.1; -.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

KW Coat protein.

FT VARIANT 416 416 D -> N.

FT VARIANT 430 430 T -> N.

FT VARIANT 540 541 QV -> HV.

FT VARIANT 573 573 V -> R.

SQ SEQUENCE 598 AA; 66619 MW; 070811ED9368E934 CRC64;

Query Match 84.9%; Score 2759.5; DB 12; Length 598;

Best Local Similarity 82.8%; Pred. No. 1.2e-191;

Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

QY 2 APGKKRPVEQSPQEPDSSSGIGKTGQQPAKRLNFGQTDSESVDPDQPLGEPATPAAV 61

Db 2 APGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTDADSVDPDQPLGQPPAAPSG 61

QY 62 GPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLY 121

Db 62 GTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLY 121

QY 122 QISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDQWRLNINNNGFRPKRLNFKLF 181

Db 122 QISSQS-GASNDNHFGYSTPWGYDFDNRFCHFSFPRDQWRLNINNNGFRPKRLNFKLF 180

QY 182 IQKVEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFMIPQYG 241

Db 181 IQKVEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYG 240

QY 242 LTLNNGSOAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 301

Db 241 LTLNNGSOAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSQSLDRLMNP 300

QY 302 IDQYLYLNRNTQNSQSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 361

Db 301 IDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 360

QY 362 SNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMGSGVMIFGKESAGASNTALDN 421

Db 361 SEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLIFGKQSGSEKTNVDIEK 420

QY 422 MITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDYLQGP 481

Db 421 MITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATAVDNTQGVLPGMVWQDRDYLQGP 480

QY 482 WAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTG 541

Db 481 WAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTG 540

QY 542 VSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599

Db 541 VSVEIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 7

Q8JQGO PRELIMINARY; PRT; 737 AA.

ID Q8JQGO

AC Q8JQGO;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Capsid protein.

OS Adeno-associated virus 7.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

OX NCBI_TaxID=202812;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=12192090;

RA Gao G.P., Alvira M.R., Wang L., Calcedo R., Johnston J., Wilson J.M.;

RA "Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy.";

RT human gene therapy.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:11854-11859(2002).

DR EMBL; AF513851; AAN03855.1; -.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

SQ SEQUENCE 737 AA; 81652 MW; D63B8CE7583140D0 CRC64;

Query Match 83.3%; Score 2709.5; DB 12; Length 737;

Best Local Similarity 82.1%; Pred. No. 7e-188;

Matches 494; Conservative 49; Mismatches 54; Indels 5; Gaps 4;

QY 1 TAPGKRPVEQSPQ-EPDSSSGIGKGTGQQPAKRLNFGQTGDSSEVPDPQLGEPPTPA 59
Db 138 TAPAKRPVEPSPQSPDSTGIGKKGQQPAKRLNFGQTGDSSEVPDPQLGEPPTAAPS 197
QY 60 AVGPTTASGGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHL 119
Db 198 SVSGTVAAGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHL 257
QY 120 YKQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLNINNWGFRPKRLNFKL 179
Db 258 YKQISSETAGSTNDNTYFGYSTPWGYFDNRFHCHFSRDPQRLNINNWGFRPKRLRFLK 317
QY 180 FNIQKVEVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQY 239
Db 318 FNIQKVEVTTNDGVTTIANNLSTIQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQY 377
QY 240 GYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMN 299
Db 378 GYLTLNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEDVPFHSSYAHQSGLDRLMN 437
QY 300 PLIDQYLYLNRTO-NQSGSAQNKOLLFSRSGSPAGMSVQPKNWLPGPCYRQQRVSKTKTD 358
Db 438 PLIDQYLYLARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQRVSKTLDQ 497
QY 359 NNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMI FGKESAGASN-TA 417
Db 498 NNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDKDFPSSGVLI FGK--TGATNKTT 555
QY 418 LDNVMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDVHAMGALPGMVWQDRDVYL 477
Db 556 LENVLTNBEIEIRPTNPVATEEYGVSSNLQAANTAAQTQVNNQOGALPGMVWQNRDVYL 615
QY 478 QGPIWAKIPHDTGHHFSPPLMGFGFLKNPPQILIKNTVPANPPAEFSATKFAFITQY 537
Db 616 QGPIWAKIPHDTGNFHPSPPLMGFGFLKHPPQILIKNTVPANPEVFTPAKFAFITQY 675
QY 538 STQVSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTR 597
Db 676 STQVSVSEIEWELQKENSKRWNPEIQYTSNFEKQTVGVDFAVDSQGVYSEPRPIGTRYLTR 735
QY 598 PL 599
Db 736 NL 737

RESULT 8
Q8JQF8 PRELIMINARY; PRT; 738 AA.
AC Q8JQF8
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Capsid protein.
OS Adeno-associated virus 8.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=202813;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12192090;
RA Gao G.P., Alvira M.R., Wang L., Calcedo R., Johnston J., Wilson J.M.;
RT "Novel adeno-associated viruses from rhesus monkeys as vectors for
human gene therapy."
RL Proc. Natl. Acad. Sci. U.S.A. 99:11854-11859(2002).
DR EMBL; AF513852; AAN03857.1; --
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 738 AA; 81756 MW; 598C2BE1B148AA14 CRC64;
Query Match 82.4%; Score 2679; DB 12; Length 738;
Best Local Similarity 80.7%; Pred. No. 1.1e-185;
Matches 485; Conservative 47; Mismatches 67; Indels 2; Gaps 2;
QY 1 TAPGKRPVEQSPQ-EPDSSSGIGKGTGQQPAKRLNFGQTGDSSEVPDPQLGEPPTPA 59

Db 138 TAPGKRPVEPSPQSPDSTGIGKKGQQPAKRLNFGQTGDSSEVPDPQLGEPPTAAPS 197
QY 60 AVGPTTASGGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHL 119
Db 198 SVSGTVAAGGAPMADNNEGADGVGNAGNWHCDSTWLGDRTVTTSTRTWALPTYNNHL 257
QY 120 YKQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLNINNWGFRPKRLNFK 178
Db 258 YKQISNGTSGATNDNTYFGYSTPWGYFDNRFHCHFSRDPQRLNINNWGFRPKRLSPK 317
QY 179 LFNIQKVEVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQ 238
Db 318 LFNIQKVEVTTNDGVTTIANNLSTIQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQ 377
QY 239 YGYLTNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLM 298
Db 378 YGYLTNNGSQAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEDVPFHSSYAHQSGLDRLM 437
QY 299 NPLIDQYLYLNRTO-NQSGSAQNKOLLFSRSGSPAGMSVQPKNWLPGPCYRQQRVSKTKTD 358
Db 438 NPLIDQYLYLSRTQTOTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTQ 497
QY 359 NNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMI FGKESAGASN-TAL 418
Db 498 NNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDKDFPSSGVLI FGKQNAARDNADY 557
QY 419 DNVMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGVDVHAMGALPGMVWQDRDVYLQ 478
Db 558 SDVMTSEEEIKTNPVATEEYGVADNLQOQNTAPQIGTVNSQOGALPGMVWQNRDVYLQ 617
QY 479 GPIWAKIPHDTGHHFSPPLMGFGFLKNPPQILIKNTVPANPPAEFSATKFAFITQYS 538
Db 618 GPIWAKIPHDTGNFHPSPPLMGFGFLKHPPQILIKNTVPADPPTTFNQSKLNSPITQYS 677
QY 539 TGQVSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRP 598
Db 678 TGQVSVSEIEWELQKENSKRWNPEIQYTSNYKSTSDFAVNTGEGVSEPRPIGTRYLTRN 737
QY 599 L 599
Db 738 L 738

RESULT 9
O92917 PRELIMINARY; PRT; 533 AA.
ID O92917
AC O92917;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE Major coat protein VP3.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
proteins affect viral infectivity: lack of an RGD integrin-binding
motif";
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03779.1; --
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.

FT VARIANT 351 351 D -> N.
FT VARIANT 365 365 T -> N.
FT VARIANT 475 476 QV -> HV.
FT VARIANT 508 508 V -> R.
SQ SEQUENCE 533 AA; 60063 MW; 9E4D8BC25810D4F0 CRC64;

Query Match 76.3%; Score 2481.5; DB 12; Length 533;
Best Local Similarity 83.3%; Pred. No. 1.5e-171;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 66 MASGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 125
Db 1 MATGSGAPMADNNEGADGVGNSSGNWHCHDSTWMGDRVTTSTRTWALPTYNNHLYKQISS 60

QY 126 ASTGASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLNNWGFPRKRLNFKLFIQVK 185
Db 61 QS-GASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLNNWGFPRKRLNFKLFIQVK 119

QY 186 EVTTNDGVTTIANLNTSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPQYGLTLN 245
Db 120 EVTQNDGTTIANLNTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGLTLN 179

QY 246 NGSQAVGRSSFCYCLEYFPSPQLRTGNFTFSYTFEEVPHSSYAHQSLSLDRMLNPLIDQY 305
Db 180 NGSQAVGRSSFCYCLEYFPSPQLRTGNFTFSYTFEEVPHSSYAHQSLSLDRMLNPLIDQY 239

QY 306 LYLNRTQNSQSAQNKKLLFRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSFT 365
Db 240 LYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNSEYS 299

QY 366 WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMIFFGKESAGASNTALDNVMTD 425
Db 300 WTGATKYHLNGRDSLVPNPGPAMASHKDDKDFPQSGVLIFGKQSEKTNVDIEKVMITD 359

QY 426 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKI 485
Db 360 EEEIRTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMWQDRDVLQGPWAKI 419

QY 486 PHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPAEFSATKFAFITQYSTGQVSVE 545
Db 420 PHTDGHFHPSPLMGGFGLKHPPQILIKNTPVPANPSTTFSAAKFAFITQYSTGQVSVE 479

QY 546 IEWELQENSKRWNPVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 599
Db 480 IEWELQENSKRWNPVQYTSNYAKSVNVDFTVDTNGVSEPRPIGTRYLTRNL 533

RESULT 10
O41855 PRELIMINARY; PRT; 734 AA.
ID O41855
AC O41855;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Capsid.
OS Adeno-associated virus 4.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=57579;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC VR-646;
RX MEDLINE=97404695; PubMed=9261407;
RA Chiorini J.A., Yang L., Liu Y., Safer B., Kotin R.M.;
RT "Cloning of adeno-associated virus type 4 (AAV4) and generation of recombinant AAV4 particles.";
RL J. Virol. 71:6823-6833 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC VR-646;
RA Chiorini J.A., Yang L., Kotin R.M., Safer B.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U89790; AAC58045.1; -.
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 734 AA; 80639 MW; 616CC27A777BBE6F CRC64;

Query Match 56.3%; Score 1830.5; DB 12; Length 734;
Best Local Similarity 57.8%; Pred. No. 3.6e-124;
Matches 355; Conservative 74; Mismatches 154; Indels 31; Gaps 10;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNF-GQTGDSSEVDPQPLGEPPATPA 59
Db 137 TAPGKKRPLIESPQPDSSSTGIGKKGQPAKRLVFEDETGAGDGPPEGSTSG-----A 190

QY 60 AVGPTTMASGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRTVTTSTRTWALPTYNNHL 119
Db 191 MSDDSEMRAAAGAAVEGGQAGDGVGNASGDWHCHDSTWSEGHVTTTSTRTWVLPPTYNNHL 250

QY 120 YKQISSASTGASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLNNWGFPRKRLNFKL 179
Db 251 YKRLGE----SLQSNITYNGFSTPWGYFDNRFCHFSPRDWQRLNNWGMRPKAMRVKI 306

QY 180 FNIQVKEVTNDGVTTIANLNTSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPQY 239
Db 307 FNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAQEGSLPFPFNDVFMVPOY 366

QY 240 GY---LTLNNGSAVGRSSFCYCLEYFPSPQLRTGNFTFSYTFEEVPHSSYAHQSLSDR 296
Db 367 GYCGLVGTNTSQQTDNRNFAFCYCLEYFPSPQLRTGNFTFSYTFEEVPHSSYAHQSLSDR 426

QY 297 LMNPLIDQYLYLNRTQN---QSGSAQNKDLLFSRSPAGMSVQPKNWLPGPCYRQQRV 352
Db 427 LMNPLIDQYLVGLQSTTTGTTLNAGTATTN---FTKLRPTNFSNFKKNWLPGPSIKQQGF 483

QY 353 SKTKTDNNNSNFTWTGAS---KY---NLNGRESIINPGTAMASHKDDKDFPMSGVMI 405
Db 484 SKTA--NQYKIPATGSDSLIKYETHSTLDGRWSALTGPPMATAGPADSK-FSNSQLIF 540

QY 406 FGKESAGASNTALDNVMTDEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGAL 465
Db 541 AGPKQNGNTATVPGLTIFTSEELAAATNATDMDMGNLPGGDQSNLPTVDRLTALGAV 600

QY 466 PGMVWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPAEF 525
Db 601 PGMVWQNRDIYQGPWAKIPHTDGHFHPSPLIIGFGLKHPPQIFIKNTPVPANPATTF 660

QY 526 SATKFASFITQYSTGQVSVEIEWELQENSKRWNPVQYTSNYAKSANVDFTVDDNGLYT 585
Db 661 SSTPVNSFITQYSTGQVSQIDWEIQKERSKRWNPVQYTSNYGQNSLLWAPDAAGKYT 720

QY 586 EPRPIGTRYLTRPL 599
Db 721 EPRAIGTRYLTRHL 734

RESULT 11
Q67666 PRELIMINARY; PRT; 732 AA.
ID Q67666
AC Q67666;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VP1.
GN VP1.
OS Goose parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=38251;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Virulent B;
RX MEDLINE=96010229; PubMed=7571426;
RA Zadori Z., Stefancsik R., Rauch T., Kisary J.;
RT "Analysis of the complete nucleotide sequences of goose and muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2.";
RL Virology 212:562-573 (1995).

DT	01-NOV-1996 (TReMBLrel. 01, Created)
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE	Capsid protein VP.
GN	VP.
OS	Barbarie duck parvovirus.
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX	NCBI_TaxID=39118;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FM;
RX	MEDLINE=96010229; PubMed=7571426;
RA	Zadori Z., Stefancsik R., Rauch T., Kisary J.;
RT	"Analysis of the complete nucleotide sequences of goose and muscovy
RT	duck parvoviruses indicates common ancestral origin with adeno-
RT	associated virus 2.";
RL	Virology 212:562-573(1995).
DR	EMBL; U22967; AAA83225.1; -.
DR	InterPro; IPR001403; Parvo_coat.
DR	Pfam; PF00740; Parvo_coat; 1.
SQ	SEQUENCE 732 AA; 81314 MW; D92360596E2D2C05 CRC64;
<div>Query Match 52.4%; Score 1702; DB 12; Length 732; Best Local Similarity 53.0%; Pred. No. 7.5e-115; Matches 325; Conservative 83; Mismatches 165; Indels 40; Gaps 11;</div>	
Qy	6 KRPVEQSPQEPDSSGIGK-TGQPAKRLNFGQTGDSSEVPDPQLG-----E 53
Db	141 EEPVNMAPAKSS----GKLTDHDPVKKPKLSE----ENSPSPNSGGESAATAEGSE 192
Qy	54 PPATPAAVGPTTMSGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITSTRTWALP 113
Db	193 PVAAP-----NMAEGGSGAMGDSAGGADGVGNASGNWHCDSQWLGDVTITKTRTWVLP 246
Qy	114 TYNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLNNNWGFRPK 173
Db	247 SYNHHMYQAITSCTNPDSN-TQYAGYSTPWGYFDNRFHCHFSPRDWQRLNNHWGIRPK 305
Qy	174 RLNFKLFNIOVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADV 233
Db	306 ALXFKIFNVQVKEVTTQDTKTIANNLTSTIQIFTDNEHQLPYVLGSATEGTMPFPFSDV 365
Qy	234 FMIPQYGYLTN---NGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEEVPFHSSYAH 290
Db	366 YALPQYGYCTMHTNQSGARFNDRSAFYCLEYFPSSQMLRTGNFFESFEFEVFPFHSMAFH 425
Qy	291 SQSLDRLMNLIDQYLYLNRNQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQ 350
Db	426 SQDLDRLMNPLLDQYLNPFSEV-NGGRNAQ-----PKKAVKGAFGAMGRNWLPGPKLLDQ 479
Qy	351 RVSKTK--TDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMGVMIFGK 408
Db	480 RVRAYSGGTDNYANWSIWSKGNKVLKDRYLLQPGVPVATHTTEDQASSVPAQNIIGIAK 539
Qy	409 E--SAGASNTALDNVMITDDEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALP 466
Db	540 DPYRSGSTLAGISDIMVTDEQEIAPTNGVGWRPPYGLTVTNEQNTTTAPTNAELEVLGALP 599
Qy	467 GMVWQDRDVYLGQPIWAKIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPANPPAEFS 526
Db	600 GMVWQNRDIYLGQPIWAKIPKTDGKPHPSPLNLGGFGLHNPPPPQVFIKNTVPADPPLEYV 659
Qy	527 ATKPASFITQYSTGQSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTE 586
Db	660 NQKWNYSYITQYSTGQCTVMVWVELRKENSKRWNPEIQFTSNFGNRTSTMFAPNETGGYVE 719
Qy	587 PRPIGTRYLTRPL 599
Db	720 DRLIGTRYLTQNL 732

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:01:27 ; Search time 13.7812 Seconds
(without alignments)
1839.046 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2764.5	85.0	735	4 US-09-321-589-1	Sequence 1, Appli
2	1830.5	56.3	598	4 US-09-532-594B-16	Sequence 16, Appl
3	1830.5	56.3	734	4 US-09-532-594B-4	Sequence 4, Appli
4	1690.5	52.0	544	4 US-09-532-594B-18	Sequence 18, Appl
5	772	23.7	756	4 US-09-438-268-4	Sequence 4, Appli
6	479.5	14.7	543	3 US-08-856-841-22	Sequence 22, Appl
7	326	10.0	415	3 US-08-856-841-20	Sequence 20, Appl
8	325	10.0	486	3 US-08-856-841-19	Sequence 19, Appl
9	325	10.0	500	3 US-08-856-841-16	Sequence 16, Appl
10	325	10.0	501	3 US-08-856-841-18	Sequence 18, Appl
11	318	9.8	395	3 US-08-856-841-13	Sequence 13, Appl
12	310	9.5	264	3 US-08-856-841-14	Sequence 14, Appl
13	303	9.3	398	3 US-08-856-841-21	Sequence 21, Appl
14	216.5	6.7	579	6 5223424-13	Patent No. 5223424
15	192	5.9	584	3 US-09-022-949-2	Sequence 2, Appli
16	178.5	5.5	387	3 US-08-856-841-17	Sequence 17, Appl
17	119	3.7	1095	4 US-09-107-532A-3855	Sequence 3855, Ap
18	115	3.5	2736	4 US-09-252-991A-30227	Sequence 30227, A
19	113.5	3.5	655	1 US-08-469-202-27	Sequence 27, Appl
20	113.5	3.5	655	2 US-08-484-434C-34	Sequence 34, Appl
21	113.5	3.5	655	4 US-09-384-361-34	Sequence 34, Appl
22	112.5	3.5	3060	2 US-08-487-826B-14	Sequence 14, Appl
23	110.5	3.4	624	3 US-08-947-965-78	Sequence 78, Appl
24	110.5	3.4	655	1 US-08-469-202-28	Sequence 28, Appl
25	110.5	3.4	655	2 US-08-484-434C-35	Sequence 35, Appl
26	110.5	3.4	655	4 US-09-384-361-35	Sequence 35, Appl
27	109	3.4	1186	1 US-08-485-568A-4	Sequence 4, Appli

28	109	3.4	1186	1 US-08-357-698-6	Sequence 6, Appli
29	109	3.4	1186	2 US-08-590-554A-4	Sequence 4, Appli
30	109	3.4	1186	2 US-09-184-223-4	Sequence 4, Appli
31	109	3.4	1186	5 PCT-US93-12682-6	Sequence 6, Appli
32	106.5	3.3	1848	3 US-08-296-791-6	Sequence 6, Appli
33	106.5	3.3	1848	5 PCT-US95-10661A-6	Sequence 6, Appli
34	104.5	3.2	717	4 US-09-626-589-1	Sequence 1, Appli
35	104	3.2	1013	3 US-09-415-522-8	Sequence 8, Appli
36	102	3.1	10182	4 US-09-134-001C-3159	Sequence 3159, Ap
37	101.5	3.1	757	4 US-09-252-991A-20231	Sequence 20231, A
38	101	3.1	1651	3 US-09-540-245A-18	Sequence 18, Appl
39	100	3.1	847	4 US-09-373-157-4	Sequence 4, Appli
40	99	3.0	769	4 US-09-252-991A-26777	Sequence 26777, A
41	99	3.0	824	4 US-09-626-589-3	Sequence 3, Appli
42	99	3.0	997	1 US-08-232-540-1	Sequence 1, Appli
43	99	3.0	997	1 US-08-428-949A-1	Sequence 1, Appli
44	99	3.0	997	1 US-08-428-948A-1	Sequence 1, Appli
45	99	3.0	997	2 US-08-428-946-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.0%; Score 2764.5; DB 4; Length 735;
Best Local Similarity 82.8%; Pred. No. 3.3e-237;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy	1	TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKRLNFGQTGDSSEVPDQPLGEPPTAA	60
Db	138	TAPGKKRPVEHSPVPDSSSGTGKAGQQPAKRLNFGQTGDADSVDPDQPLGQPPA	197
Qy	61	VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITSTRTWALPT	120
Db	198	LGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITSTRTWALPT	257
Qy	121	KQISSASTGASNDNHFGYSTPWGYFDENRPHCHFPSRDWQRLNNNWGFRPKRL	180
Db	258	KQISSQS-GASNDNHFGYSTPWGYFDENRPHCHFPSRDWQRLNNNWGFRPKRL	316
Qy	181	NIQKVEVTTNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQGLCPFPFADVF	240
Db	317	NIQKVEVTTNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQGLCPFPFADVF	376
Qy	241	YLTNNGSQAVGRSSFYCLEYFPPSOMLTGNNFTFSYTFEEVPPHSSVAHSQSL	300
Db	377	YLTNNGSQAVGRSSFYCLEYFPPSOMLTGNNFTFSYTFEDVPFHSSVAHSQSL	436
Qy	301	LIDQLYLNLRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYQORVSKTKTD	360
Db	437	LIDQLYLNLRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYQORVSKTKTD	496
Qy	361	NSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFPMGVMIFGKESAGAS	420
Db	497	NSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKDFPFQSGVLI FGQSEKTN	556

Qy	406	FGKESAGASNTALDNVMTDEEEIKATNPVATERFGTVAVNFOSSSTDTPATGDVHVAMGAL	465
Db	541	AGPKQNGNTATVPGTLTFTSEEEAATNATDTDMWGNLPGGDQSNLNLP TVDRLTALGAV	600
Qy	466	PGMVQDRD VYLQGP I WAKI PHTDGHFHPSP LMGFGFGLKGNPPQ I L I KNTVPVPANPPAEF	525
Db	601	PGMVQNRDIYYQGP I WAKI PHTDGHFHPSP LIGGFGFGLKHPPPQ I F I KNTVPVPANPATTF	660
Qy	526	SATKPASFITQYSTGQVSVEIEWELQKENS KRWNPVEQYTSNYAKSANVDFTVDNNGLYT	585
Db	661	SSTPVNSFITQYSTGQVSQVIDWEIQKERS KRWNPVEQFTSNYGQNSLLWAPDAAGKYT	720
Qy	586	EPRPIGTRYLTRPL	599
Db	721	EPRAIGTRYLTHHL	734

```

RESULT 4
US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18

```

Query Match	52.0%;	Score 1690.5;	DB 4;	Length 544;
Best Local Similarity	59.4%;	Pred. No. 7.5e-142;		
Matches 325;	Conservative	65;	Mismatches 132;	Indels 25;
				Gaps 9;

Qy	67	ASGGGAPMADNNEGADGVGNASGNWHCDSTWLGD	RVTITSTRTWALPTYNNHLYKQISSA	126
Db	9	AAAGGA	AV-EGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLP	66
Qy	127	STGASNDNH	FGYSTPGWGFDFNRFCHFS	186
Db	67	---	SLQSN	123
Qy	187	VTNDGVTTI	ANNLTSTVQVFS	243
Db	124	VTTSG	ETTVA	183
Qy	244	LNNGSQ	AVGRSSFYCLEYF	303
Db	184	GNTSQ	QQTDRNAFYCLEYF	243
Qy	304	QYLYLNR	TQN---QSGSAQNKDLLFSR	359
Db	244	QYLWGL	QSTTTTGTTLNAGTATTN---	298
Qy	360	NNSNFT	WTGAS---KY---NLNGRESI	412
Db	299	QNYKI	PATGSDSLIKYETHSTLDGRWSAL	357
Qy	413	ASNTALD	NVMITDEEIEIKATNPVATER	472

Qy	398	FPMGVMIFGKESAGASNTALDNV-----MITDEEEIKATNPVATERFGTVAVNFQSS	450
Dd	548	VTGINAISHGQTTYG---NAEDKEYQGVRFPNEKEQLKLGLNMHTY-----FPNK	598
Qy	451	STDPATGDVHAMGALPGMVWQRDRDYYLQQPIWAKI PHTDGHFHPS-PLMGSGFLKNPPPO	509
Dd	599	GTOYTDDQIE-RPLMVGSVNRRALHYESQLWSKIPNLDDSFKTQFAALCGWLHQPPPO	657
Qy	510	ILIKTTPVPANPPAEFSATKFASFITQYSTGOVSVEIEWEL-QKENSKRWNPE	561
Dd	658	IPLK--ILPQSGPIGGIKSMGITTLVQYAVGIMTVMTFKLGPRKATGWNPQ	708

RESULT 6
US-08-856-841-22
; Sequence 22, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:

IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match	14.7%;	Score 479.5;	DB 3;	Length 543;
Best Local Similarity	27.1%;	Pred. NO. 4.2e-34;		
Matches 141; Conservative	79;	Mismatches 226;	Indels 75;	Gaps 16
Qy	75	ADNNEGADGVGNAS--GNWHCDSTWLGDVRVITSTRTWTALPTYNHLYKQISSASTGASN	132	
		: : : : : : : : :		
Db	16	AEASTGAGCGGSNVKSMWSEGATFSANSVCTFSRQFLIPYDPEHHYKVFSPPAASSCHN	75	
Qy	133	D-----NHYFGYSTPWGYDFDNRFHCHFSPRDQQLINNNWGFRLNFKLFNI	182	
		: : : : : : :		
Db	76	ASGKEAKVCTISPIMGYSTPWRYLDFNALNLFPSLEFQHLIENYGSAPDALTVTISEI	135	

RESULT 7
US-08-856-841-20
; Sequence 20, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
;

STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856.841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214.658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917.096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 10.0%; Score 326; DB 3; Length 415;
Best Local Similarity 26.0%; Pred. No. 1.3e-20;
Matches 101; Conservative 66; Mismatches 180; Indels 42; Gaps 13;
QY 193 VTTIANNLTSTVQFSDSEYQLPYVLGSAHQGLPPFPADVPMIPQYGLTLNN-GSQAV 251

Db 1 MTMITPSLHACMLV--DHEYKYPYVLGGQDQLAPLPIWVYPPQYAYLTVGDVNTQGI 58
QY 252 G-----RSSPYCLEYFPQSMLRTGNFTFSYTFEEVPHSSYAHQSLSLRLMNP LI 302
Db 59 SGDSKCLASEESAFYVLEHSSFQLLGTGTASMSYKFPVPPELEGCSQHFYEMYNPL- 117
QY 303 DQYLYLNRTQNSGAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNS 362
Db 118 --YGSRLGVDPDTLGGDPKFRSL-----THEDHAIQPNFMPGPLVNSVSTKEGDSNTGA 170
QY 363 NFTWTGASKYNLNGRESIINPG-TAMASHKDDDDKFFPMGVMIFGKESAGASNTALDNV 421
Db 171 GKALTGLSTGTSQNTTRISLRPGVSPQYHHWTDKVVYTG--NAEDKE 227
QY 422 -----MITDEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGVMVQD RD 474
Db 228 YQGVGRFPNEKEQLKQLGLNMHTY-----FPNKGTYQYTDQIE-RPLMVGSVWNRRA 280
QY 475 VYLGPIWAKIPHTDGHFHPs-PLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFA SP 533
Db 281 LHYESQLWSKIPNLDDSFKTQFAALGGWGLHQPFPQIFLK--ILPESGPIGGIKSMGITT 338
QY 534 ITQYSTGQVSVEIEWEL-QKENSQRWNP 561
Db 339 LVQYAVGIMTVTMTFKLGP RKATGRWNPQ 367
RESULT 8
US-08-856-841-19
; Sequence 19, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856.841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214.658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917.096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:

Db 319 LKQYAVGIMTVTMTFKLGPRKATGRWNPO 347

RESULT 12

```

US-08-856-841-14
; Sequence 14, Application US/08856841
; Patent No. 6274307
;
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
;
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LXR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
;
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
;
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
; IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
; POSITION IN GENOME: N/A
; FEATURE:
;
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
;
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; AUTHORS: WIDDOWS, D.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET

```


Db 117 WGVWFNPADWQLISNNMTINLVSFEQEIFNVVLKTTITESATSPPSKIYNNDLTASLMVA 176
QY 208 SDSEYQLPYVLGSAHQCLPPFPADVFMIPQGYL-----TLNNGSQAAGRS--- 254
Db 177 LDTNNTLPYTAPRSETLGPYPWLPTKPTQRYLYLSCLRNLPPTYTQSQQITDSIQT 236
QY 255 -----SFYCLE-YFPSQMLRTGNMF-TFSYTFEEVPHSSYAHSQSLDRLMNLIDQYL 306
Db 237 GLHSDIMFYTIENAVPIHLRTGDEFSTGIYHFDTKPL--KLTHSWQTNR----- 284
QY 307 YLNRQTQNSGSAQNKDLLFSRGSAGMSVQP-----KNWLPGPCYRQ---QRVSKTK 356
Db 285 -----SLGLPPKVLTEPTTEGDOHPGTLPGANTRKGYHQTINNSY 324
QY 357 TD-----NNNSNFTWTGASKYNLNGRESIINP--GTAMASHKDDDEKFPMSG 402
Db 325 TEATALRPAQVGYNTPYMFDYSGGPF-----LTPIVPTADTQYYDDE----PNGA 372
QY 403 VMIFGKESAGASNTALDNVMTDEEBEIKATNP-----VATERFGTVA-VNFQSSS---- 451
Db 373 IRTMGYQHGLTTS-----SOELERYTFNPQSKGRAPKQQFNQQAPLNLENTNNGTL 426
QY 452 --TDPATG--DVHAMGALP-----GMVWQDRDVYLOGPIWAKIPHTD--GHFHPSPLMGG 500
Db 427 LPSDPGKSNKHFMNTLNTYGLPTALNNTAPVFPNGQIWDKELDTDLKPRLH---VTAP 483
QY 501 FGLK-NPPPPQILIKNTVPANPAEESA--TKFASFITYSTGVSVEIEWELQKENSKR 557
Db 484 FVCKNPPGQLEVKIAP---NLTDDEFNADSPQQPRIITD-SNFWWKGTLTFTAKMRSSNM 539
QY 558 WNPEVOYTSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 540 WNPLOQHTT-----TAENIRKYI-PTNIG 562

RESULT 15

US-09-022-949-2
; Sequence 2, Application US/090222949
; Patent No. 6187759
; GENERAL INFORMATION:
; APPLICANT: Tarpey, Ian
; APPLICANT: Greenwood, Neil
; TITLE OF INVENTION: Canine parvovirus DNA vaccination
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 6187759el Patent Dept.
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,949
; FILING DATE: 11-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-022-949-2
Query Match 5.9%; Score 192; DB 3; Length 584;
Best Local Similarity 19.9%; Pred. No. 1.8e-08;
Matches 128; Conservative 100; Mismatches 232; Indels 184; Gaps 32;
QY 49 QPLGEPPA--TPAAVGPPTTMASGGGAPMADNNEGADGVGNASGNWH-----CDSTWLG 99
Db 7 QPDGGQPAVRNERATSGNGSGGG-----GGSGGVGISTGTFFNNQTEPKFLENGWV- 59
QY 100 DRVITTSRTIWAALPTNNHLYKQI-----SSASTGASNDNHYFGYSTPWGYFDFNRFH 152
Db 60 -EITANSSRLVHLNMPESENYRRVVVNNLDKTAVNGNMALDDTHAQIVTPWSLVDANAWG 118
QY 153 CHFSPRDWQRLINNNWGFPRKRLNFKLFNIQKVT---TNDGVTTIANNLSTVQVPSD 209
Db 119 WFNPGDWQLLVNMTSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALD 178
QY 210 SEYQLPYVLGSAHQGCL-----PPPPA-----DVFMIPQGYLTLNNGSQAAGRS- 254
Db 179 SNNTMPFTPAAMRSETLGFYPWKPTIPTPWRYYFQWDRTLIPSH-----TGTSGTPTNI 232
QY 255 -----SFYCLE-YFPSQMLRTGNMF-TFSYTFEEVPHSSYAHSQSLDRLMNLID 303
Db 233 YHGTDPODDVQFYTIENSVPVHLLRTGDEFATGTFFFDCKP--CRLTHTWQTNRALG--LP 288
QY 304 QYLYLNRQTQNS-----GSAQNK-----DLLFSRGSAGMSVQPKNWLP 344
Db 289 PFLNSLPQSEGGTFNGYIGVQODKRRGVQTMGNTNYITEATIMRPAEVGYS-----A 340
QY 345 PCYRQORVSK-----TKTDNNSNFTWTGASKY---NLNGRESIINPGT--- 385
Db 341 PYSFEASTQGPFTPIAAGRGAQTDEQA---DGDPRYAFGRQHGOKTTTGETPER 397
QY 386 -AMASHKDDDEKFPMSGVMI FGKESAGASNTALD-NVMITDEEEIKATNPVATERFGTV 443
Db 398 FTYIAHQDT-----GRYPEGDWTONINFNLPVTNDNVLLPTDPIG-----GKT 440
QY 444 AVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVYLOGPIWAKIPHTDGHFHPSPLMGGFGL 503
Db 441 GINY--TNIFNTYGLTALNNVP-----PVYPNGQIWDKEFTD-----L 478
QY 504 K-----NPPPOILIKNTVPAN---PPAEFSATKFASFITQYSTGQVSVEIE 547
Db 479 KPRLHVNAPFVCQNNCPGQLFVKVAPNLNTNEYDPPDASANMSR-----IVTYSDFWWMKGKLV 534
QY 548 WELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 591
Db 535 FKAKLRASHTWNPIQQMSIN-----IDNQFNYYV-PSNIG 567

Search completed: January 21, 2004, 16:08:02
Job time : 15.7812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:04:27 ; Search time 26.9213 Seconds
(without alignments)
4549.706 Million cell updates/sec

Title: US-09-807-802A-15
Perfect score: 3251
Sequence: 1 TAPGKKRPVEQSPQEPDSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3251	100.0	736	12	US-10-291-583-64	Sequence 64, Appl
2	3251	100.0	736	12	US-10-423-704A-5	Sequence 5, Appli
3	3229	99.3	736	12	US-10-291-583-65	Sequence 65, Appl
4	2815	86.6	736	12	US-10-291-583-71	Sequence 71, Appl
5	2815	86.6	736	12	US-10-423-704A-6	Sequence 6, Appli
6	2764.5	85.0	735	12	US-10-291-583-70	Sequence 70, Appl
7	2764.5	85.0	735	12	US-10-423-704A-4	Sequence 4, Appli
8	2764.5	85.0	735	14	US-10-038-972A-13	Sequence 13, Appl
9	2764.5	85.0	735	15	US-10-293-478-1	Sequence 1, Appli
10	2759.5	84.9	598	14	US-10-038-972A-14	Sequence 14, Appl
11	2753.5	84.7	735	12	US-10-291-583-67	Sequence 67, Appl
12	2751.5	84.6	735	12	US-10-291-583-68	Sequence 68, Appl
13	2749.5	84.6	735	12	US-10-291-583-69	Sequence 69, Appl
14	2745.5	84.5	735	12	US-10-291-583-66	Sequence 66, Appl
15	2735	84.1	736	12	US-10-291-583-97	Sequence 97, Appl

ALIGNMENTS

RESULT 1

```

US-10-291-583-64
; Sequence 64, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting
; TITLE OF INVENTION: Sequences and Isolates
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype
US-10-291-583-64

```

Query Match	100.0%;	Score 3251;	DB 12;	Length 736;
Best Local Similarity	100.0%;	Pred. NO. 1.2e-283;		
Matches 599;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDPQPLGPEPPATPAA	60	
Db	138	TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDPQPLGPEPPATPAA	197	
Oy	61	VGPTTMASSGGCAPMADNNNEGADGVGNASGNWHCDSTWLGDRTVITTTTWTALPTVYNNHLY	120	

Db 198 VGPTTASGGGAPMADNNEGADGVGNASGNHWCDSWLGDVRVTTSTRTWALPTYNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDWQRLINNNWGFPRKRLNFKLF 180
Db 258 KOISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDWQRLINNNWGFPRKRLNFKLF 317
Qy 181 NIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 240
Db 318 NIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 377
Qy 241 YLTLNNGSAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNP 300
Db 378 YLTLNNGSAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNP 437
Qy 301 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 438 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGVMIFGKESAGASNTALDN 420
Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGVMIFGKESAGASNTALDN 557
Qy 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGP 480
Db 558 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGP 617
Qy 481 IWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQILIKNTVPVPPANPPAEFSATKFASFTQYSTG 540
Db 618 IWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQILIKNTVPVPPANPPAEFSATKFASFTQYSTG 677
Qy 541 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 599
Db 678 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 736

RESULT 2
US-10-423-704A-5
; Sequence 5, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-423-704A-5

Query Match 100.0%; Score 3251; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.2e-283;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSESVDPQPLGEPPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSESVDPQPLGEPPATPAA 197

Qy 61 VGPTTASGGGAPMADNNEGADGVGNASGNHWCDSWLGDVRVTTSTRTWALPTYNHLY 120
Db 198 VGPTTASGGGAPMADNNEGADGVGNASGNHWCDSWLGDVRVTTSTRTWALPTYNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDWQRLINNNWGFPRKRLNFKLF 180
Db 258 KOISSASTGASNDNHFGYSTPWGYDFDNRFCHFSFPRDWQRLINNNWGFPRKRLNFKLF 317
Qy 181 NIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 240
Db 318 NIQKVEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 377
Qy 241 YLTLNNGSAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNP 300
Db 378 YLTLNNGSAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNP 437
Qy 301 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 438 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 497
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGVMIFGKESAGASNTALDN 420
Db 498 NSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGVMIFGKESAGASNTALDN 557
Qy 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGP 480
Db 558 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGP 617
Qy 481 IWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQILIKNTVPVPPANPPAEFSATKFASFTQYSTG 540
Db 618 IWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQILIKNTVPVPPANPPAEFSATKFASFTQYSTG 677
Qy 541 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 599
Db 678 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 736

RESULT 3
US-10-291-583-65
; Sequence 65, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (A
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV6VP1
US-10-291-583-65

Query Match 99.3%; Score 3229; DB 12; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.1e-281;
Matches 594; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDQPLGEPPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSSEVPDQPLGEPPATPAA 197

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKGTQQPAKRLNFGQTGDSSESVDPQPLGEPATPAA 60
Db 138 TAPGKKRPVHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQPPAAPSG 197
QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNHWHCDSTWLGDRVITTTSTRTWALPTYNNHLY 120
Db 198 LGTNTMATGSGAPMADNNEGADGVGNSSGNHWHCDSTWMDRVITTTSTRTWALPTYNNHLY 257
QY 121 KQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 180
Db 258 KQISSQS-GASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 316
QY 181 NIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMPQYG 240
Db 317 NIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMPQYG 376
QY 241 YLTLNNGSQAVGRSSFCYCLEYFPPSQMLRTGNFTFSYTFEEVPFHSSYAHQSQSLDRLMNP 300
Db 377 YLTLNNGSQAVGRSSFCYCLEYFPPSQMLRTGNFTFSYTFEDVPFHSSYAHQSQSLDRLMNP 436
QY 301 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 437 LIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496
QY 361 NSNFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGVMIFGKESAGASNTALDN 420
Db 497 NSEYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFFPQSGVLIFGKQSGSEKTNVDIEK 556
QY 421 VMITDEEEIKATNPVATERFGTVAVNFOSSSTD PATGDVHAMGALPGMVWQDRDVLQGP 480
Db 557 VMITDEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGP 616
QY 481 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPPANPPAEFSATKFASFITQYSTG 540
Db 617 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPPANPSTTFSAAKFASFITQYSTG 676
QY 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 599
Db 677 QVSVEIEWELQKENSKRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 8

US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP1 capsid protien
US-10-038-972A-13

Query Match 85.0%; Score 2764.5; DB 14; Length 735;
Best Local Similarity 82.8%; Pred. No. 7.8e-240;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKGTQQPAKRLNFGQTGDSSESVDPQPLGEPATPAA 60
Db 138 TAPGKKRPVHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQPPAAPSG 197
QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNHWHCDSTWLGDRVITTTSTRTWALPTYNNHLY 120
Db 198 LGTNTMATGSGAPMADNNEGADGVGNSSGNHWHCDSTWMDRVITTTSTRTWALPTYNNHLY 257

QY 121 KQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 180
Db 258 KQISSQS-GASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 316
QY 181 NIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMPQYG 240
Db 317 NIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMPQYG 376
QY 241 YLTLNNGSQAVGRSSFCYCLEYFPPSQMLRTGNFTFSYTFEEVPFHSSYAHQSQSLDRLMNP 300
Db 377 YLTLNNGSQAVGRSSFCYCLEYFPPSQMLRTGNFTFSYTFEDVPFHSSYAHQSQSLDRLMNP 436
QY 301 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360
Db 437 LIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496
QY 361 NSNFTWTGASKYNLNGRESINPGTAMASHKDDDEKFFPMSGVMIFGKESAGASNTALDN 420
Db 497 NSEYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFFPQSGVLIFGKQSGSEKTNVDIEK 556
QY 421 VMITDEEEIKATNPVATERFGTVAVNFOSSSTD PATGDVHAMGALPGMVWQDRDVLQGP 480
Db 557 VMITDEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGP 616
QY 481 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPPANPPAEFSATKFASFITQYSTG 540
Db 617 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPPANPSTTFSAAKFASFITQYSTG 676
QY 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 599
Db 677 QVSVEIEWELQKENSKRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 9

US-10-293-478-1
; Sequence 1, Application US/10293478
; Publication No. US20030078411A1
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 85.0%; Score 2764.5; DB 15; Length 735;
Best Local Similarity 82.8%; Pred. No. 7.8e-240;
Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

QY 1 TAPGKKRPVEQSPQEPDSSSGIGKGTQQPAKRLNFGQTGDSSESVDPQPLGEPATPAA 60
Db 138 TAPGKKRPVHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQPPAAPSG 197
QY 61 VGPTTMASGGGAPMADNNEGADGVGNASGNHWHCDSTWLGDRVITTTSTRTWALPTYNNHLY 120
Db 198 LGTNTMATGSGAPMADNNEGADGVGNSSGNHWHCDSTWMDRVITTTSTRTWALPTYNNHLY 257
QY 121 KQISSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 180
Db 258 KQISSQS-GASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNNWGFPRKRLNFKLF 316
QY 181 NIQVKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMPQYG 240

Db 317 NIQKVEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYG 376

Qy 241 YLTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEEVPHSSYAHSQSLDRLMNP 300

Db 377 YLTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEDVPHSSYAHSQSLDRLMNP 436

Qy 301 LIDQYLYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360

Db 437 LIDQYLYLSRTNTPSGTTTQSRQLFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496

Qy 361 NSNFTWTGASKYNLNGRESINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDN 420

Db 497 NSEYSWTGATKYHLNGRDSLVPNPGPAMASHKODEEKFFPQSGVLIFGKQGSEKTNVDIEK 556

Qy 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGP 480

Db 557 VMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGP 616

Qy 481 IWAKIPHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPVPANPPAEFSATKPFASFITQYSTG 540

Db 617 IWAKIPHTDGHFHPSPMLMGFGFLKHPPPPQILIKNTVPVPANPSTTFSAAKPFASFITQYSTG 676

Qy 541 QVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599

Db 677 QVSVEIEWELQKENSKRWNPEIQYTSNYKSNVNDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 10

US-10-038-972A-14

; Sequence 14, Application US/10038972A

; Publication No. US20020192823A1

; GENERAL INFORMATION:

; APPLICANT: J. Bartlett

; TITLE OF INVENTION: AAV VECTORS AND METHODS

; FILE REFERENCE: 28335/36996US

; CURRENT APPLICATION NUMBER: US/10/038,972A

; CURRENT FILING DATE: 2002-01-04

; PRIOR APPLICATION NUMBER: US 60/260,124

; PRIOR FILING DATE: 2001-01-05

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 14

; LENGTH: 598

; TYPE: PRT

; ORGANISM: adeno-associated virus 2 VP2 capsid protien

US-10-038-972A-14

Query Match 84.9%; Score 2759.5; DB 14; Length 598;

Best Local Similarity 82.8%; Pred. No. 1.6e-239;

Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;

Qy 2 APGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDPQPLGEPPTAAV 61

Db 2 APGKKRPVEHSPVEPDSSSGTGAGQOPARKRLNFGQTGDADSVDPDPQLGQPPAAPSG 61

Qy 62 GPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYK 121

Db 62 GTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITSTRTWALPTYNNHLYK 121

Qy 122 QISSASTGASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKRLNFKLFN 181

Db 122 QISSQS-GASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKRLNFKLFN 180

Qy 182 IQVKEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPQYG 241

Db 181 IQVKEVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYG 240

Qy 242 LTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEEVPHSSYAHSQSLDRLMNP 301

Db 241 LTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEDVPHSSYAHSQSLDRLMNP 300

Qy 302 IDQYLYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 361

Db 301 IDQYLYLSRTNTPSGTTTQSRQLFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 360

Qy 362 SNFTWTGASKYNLNGRESINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNV 421

Db 361 SEYSWTGATKYHLNGRDSLVPNPGPAMASHKODEEKFFPQSGVLIFGKQGSEKTNVDIEK 420

Qy 422 MITDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGP 481

Db 421 MITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGP 480

Qy 482 WAKIPHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPVPANPPAEFSATKPFASFITQYSTG 541

Db 481 WAKIPHTDGHFHPSPMLMGFGFLKHPPPPQILIKNTVPVPANPSTTFSAAKPFASFITQYSTG 540

Qy 542 VSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599

Db 541 VSVEIEWELQKENSKRWNPEIQYTSNYKSNVNDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 11

US-10-291-583-67

; Sequence 67, Application US/10291583

; Publication No. US20030138772A1

; GENERAL INFORMATION:

; APPLICANT: Gao, Guangping

; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio

; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (/

; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1

; FILE REFERENCE: UPN-02735USA

; CURRENT APPLICATION NUMBER: US/10/291,583

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/377,066

; PRIOR FILING DATE: 2002-05-01

; PRIOR APPLICATION NUMBER: US 60/386,675

; PRIOR FILING DATE: 2002-06-05

; NUMBER OF SEQ ID NOS: 120

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 67

; LENGTH: 735

; TYPE: PRT

; ORGANISM: capsid protein of AAV serotype, clone A3.7

US-10-291-583-67

Query Match 84.7%; Score 2753.5; DB 12; Length 735;

Best Local Similarity 82.8%; Pred. No. 7.7e-239;

Matches 496; Conservative 42; Mismatches 60; Indels 1; Gaps 1;

Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDSSEVPDPQPLGEPPTAA 60

Db 138 TAPGKKRPIEQSPAEPDSSSGIGKSGQQPAKKRLNFGQTGDTESVPDPQPIGEPAPA 197

Qy 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLY 120

Db 198 VGSNTMASGGGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITSTRTWALPTYNNRLY 257

Qy 121 KOISSASTGASNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKRLNFKLF 180

Db 258 KOISSES-GATNDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKLNFKLF 316

Qy 181 NIQKVEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPQYG 240

Db 317 NIQKVEVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYG 376

Qy 241 YLTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEEVPHSSYAHSQSLDRLMNP 300

Db 377 YLTLNNGSQAVGRSSFYCLEYFPQSOMLRGTGNNFTSYTFEDVPHSSYAHSQSLDRLMNP 436

Qy 301 LIDQYLYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNN 360

Db 437 LIDQLYYLSKTQGTQSRLOFSQAGPSSMAQQAQKWLFGPSYRQORMSKTANDNN 496
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMSGVMI FGKESAGASNTALDN 420
Db 497 NSEFAWTAATKYIYLNGRNSLVNPGPPMASHKDDDEKYFPMHGNLIFGKQGTGTTNVDI ES 556
Qy 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRD VYLQGP 480
Db 557 VLITDEEEIRTTNPVATEQYGVATNHQSNTTASYSVDSQGILPGMWQDRD VYLQGP 616
Qy 481 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPANPPAEFSATKFPASFITQYSTG 540
Db 617 IWAKTPHTDGHFHPSPLMGGFGLKHPPPOILIKNTVPVPANPATTTFTPGKFPASFITQYSTG 676
Qy 541 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 677 QVSVEIEWELQKENS KRWNPEIQYTSNYKSVNVEFTVDANGVYSEPRPIGTRYLTRNL 735

RESULT 12
US-10-291-583-68
; Sequence 68, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.4
US-10-291-583-68

Query Match 84.6%; Score 2751.5; DB 12; Length 735;
Best Local Similarity 82.6%; Pred. No. 1.2e-238;
Matches 495; Conservative 43; Mismatches 60; Indels 1; Gaps 1;
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKRLNFGQTGDSSEVPDPQPLGEPPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGESGQQPAKRLNFGQTGDTESVPDPQPIGEPPAAPS 197
Qy 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVITSTRTWALPTNNHLY 120
Db 198 VGSNTMASGGGAPMADNNEGADGVGNSSGNWCHDSTWMDRVITSTRTWALPTNNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLINNNGWFRPKLNFKLF 180
Db 258 KQISSSES-GATNDNHFGYSTPWGYDFNRFCHFSPRDWQRLINNNGWFRPKLNFKLF 316
Qy 181 NIOQKEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 240
Db 317 NIOQKEVTNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 376
Qy 241 YLTLNNGSQAVGRSSFCYCLEYFPQSOMLRTGNNTFSTFTFEEVPFHSSYAHSQSLDRLMNP 300
Db 377 YLTLNNGSQAVGRSSFCYCLEYFPQSOMLRTGNNTFSTFTFEDVPFHSSYAHSQSLDRLMNP 436

Qy 301 LIDQLYYLNRTQNSGSAQNKDILLFSRGSAGMSVQPKMWLPGPCYRQORVSKTKTDNN 360
Db 437 LIDQLYYLSKTQGTQSRLOFSQAGPSSMAQQAQKWLFGPSYRQORMSKTANDNN 496
Qy 361 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMSGVMI FGKESAGASNTALDN 420
Db 497 NSEFAWTAATKYIYLNGRNSLVNPGPPMASHKDDDEKYFPMHGNLIFGKQGTGTTNVDI ES 556
Qy 421 VMITDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRD VYLQGP 480
Db 557 VLITDEEEIRTTNPVATEQYGVATNHQSODTTASYSVDSQGILPGMWQDRD VYLQGP 616
Qy 481 IWAKIPHTDGHFHPSPLMGGFGLKNPPPOILIKNTVPVPANPPAEFSATKFPASFITQYSTG 540
Db 617 IWAKTPHTDGHFHPSPLMGGFGLKHPPPOILIKNTVPVPANPATTTFTPGKFPASFITQYSTG 676
Qy 541 QVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 599
Db 677 QVSVEIEWELQKENS KRWNPEIQYTSNYKSVNVEFTVDANGVYSEPRPIGTRYLTRNL 735

RESULT 13
US-10-291-583-69
; Sequence 69, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.5
US-10-291-583-69

Query Match 84.6%; Score 2749.5; DB 12; Length 735;
Best Local Similarity 82.6%; Pred. No. 1.8e-238;
Matches 495; Conservative 43; Mismatches 60; Indels 1; Gaps 1;
Qy 1 TAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKRLNFGQTGDSSEVPDPQPLGEPPATPAA 60
Db 138 TAPGKKRPVEQSPQEPDSSSGIGESGQQPAKRLNFGQTGDTESVPDPQPIGEPPAAPS 197
Qy 61 VGPTTMASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVITSTRTWALPTNNHLY 120
Db 198 VGSNTMASGGGAPMADNNEGADGVGNSSGNWCHDSTWMDRVITSTRTWALPTNNHLY 257
Qy 121 KOISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLINNNGWFRPKLNFKLF 180
Db 258 KQISSSES-GATNDNHFGYSTPWGYDFNRFCHFSPRDWQRLINNNGWFRPKLNFKLF 316
Qy 181 NIOQKEVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 240
Db 317 NIOQKEVTNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPOYG 376
Qy 241 YLTLNNGSQAVGRSSFCYCLEYFPQSOMLRTGNNTFSTFTFEEVPFHSSYAHSQSLDRLMNP 300
Db 377 YLTLNNGSQAVGRSSFCYCLEYFPQSOMLRTGNNTFSTFTFEDVPFHSSYAHSQSLDRLMNP 436

Db

378

GXLTLNNGSQALGRSSFYCLEYFPFSQMLRTGNNFQSYTFEDVPFHSSYAHSQSLDRLMN

437

QY

300

PLIDQYLYLNRTONQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN

359

Db

438

PLIDQYLYYLVRTQT-TGTGGTQTLAFSQAGPSSMANQARNWVPGPCYRQQRVSTTTNQN

496

QY

360

NNSNFTWTGASKYNLNGRESINPGTAMASHKODEDKFPPMSGVMIFGKESAGASNTALD

419

Db

497

NNSNFAWTGAAKFKLNGRDSLNNPGVAMASHKODDDRRFPSSGVLIFGKQAGNDGVDYS

556

QY

420

NVMTDDEEEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAMGALPGMVWQDRDVYLQG

479

Db

557

QVLITDEEEIKATNPVATEEYGAVAINNQAANTQAQTGLVHNQGVIPGMVWQNRDVYLQG

616

QY

480

PIWAKIPHTDGHFHPSPMLMGFGGLKNPPPPQILIKNTPVPANPPAEFSATKFASFITQYST

539

Db

617

PIWAKIPHTDGNFHPSPMLMGFGGLKHPPPPQILIKNTPVPADPPPLTFNQAKLNSFITQYST

676

QY

540

GQVSVEIEWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

599

Db

677

GQVSVEIEWELOKENSKRWNPEIQYTSNYKSTNVDFAVNTEGVYSEPRPIGTRYLTRNL

736

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 15:59:46 ; Search time 32.5714 Seconds
(without alignments)
2602.280 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	534	21 AAY71169	Adeno-associated v
2	2906	100.0	599	21 AAY71168	Adeno-associated v
3	2906	100.0	736	21 AAY71167	Adeno-associated v
4	2884	99.2	736	22 AAB59847	AAV6 capsid protei
5	2530	87.1	736	22 AAB59846	AAV3B capsid prote
6	2528	87.0	736	22 AAB59845	AAV3A capsid prote
7	2481.5	85.4	533	22 AAG65794	Adeno-associated v
8	2481.5	85.4	533	23 AAU98976	Adeno-associated v
9	2481.5	85.4	598	22 AAG65793	Adeno-associated v

10	2481.5	85.4	598	22 AAM51509	Adeno-associated v
11	2481.5	85.4	598	23 AAU98975	Adeno-associated v
12	2481.5	85.4	735	22 AAG65792	Adeno-associated v
13	2481.5	85.4	735	22 AAM51508	Adeno-associated v
14	2481.5	85.4	735	22 AAB59844	AAV2 capsid protei
15	2481.5	85.4	735	23 AAU98974	Adeno-associated v
16	2477.5	85.3	533	22 AAM51510	Adeno-associated v
17	2455	84.5	734	22 AAB50326	Adeno-associated v
18	1690.5	58.2	544	19 AAW46314	AAV4 VP3 coat prot
19	1690.5	58.2	544	24 ABG73941	Adeno-associated v
20	1690.5	58.2	598	19 AAW46313	AAV4 VP2 coat prot
21	1690.5	58.2	598	24 ABG73940	Adeno-associated v
22	1690.5	58.2	734	19 AAW46308	AAV4 VP1 capsid pr
23	1690.5	58.2	734	24 ABG73935	Adeno-associated v
24	1690	58.2	736	18 AAW19000	Adeno-associated v
25	1676	57.7	534	16 AAR85386	Duck parvovirus ca
26	1676	57.7	732	16 AAR85385	Barbary duck parvo
27	1665	57.3	532	21 AAY58162	Barbary duck parvo
28	1665	57.3	532	23 AAU11407	Adeno-associated v
29	1665	57.3	588	21 AAY58161	Adeno-associated v
30	1665	57.3	588	23 AAU11406	Adeno-associated v
31	1665	57.3	724	21 AAY58160	Adeno-associated v
32	1665	57.3	724	23 AAU11405	Adeno-associated v
33	491.5	16.9	756	21 AAY71231	Capsid protein enc
34	490	16.9	554	16 AAW08987	Human parvovirus V
35	490	16.9	554	20 AAY23230	Erythrovirus V9 VP
36	490	16.9	781	16 AAW08986	Human parvovirus V
37	490	16.9	781	20 AAY23227	Erythrovirus V9 VP
38	486	16.7	554	24 ABP57264	Human parvovirus B
39	486	16.7	554	24 ABP57267	Human parvovirus B
40	486	16.7	781	24 ABP57263	Human parvovirus B
41	477	16.4	781	24 ABP57266	Human parvovirus B
42	476.5	16.4	543	12 AAR13405	Parvo virus B19 VP
43	304.5	10.5	264	12 AAR13407	Parvo virus B19 PA
44	295.5	10.2	370	12 AAR13406	Parvo virus B19 PA
45	234.5	8.1	579	13 AAR29079	Porcine Parvovirus

ALIGNMENTS

RESULT 1

AA71169
ID AAY71169 standard; Protein; 534 AA.

XX
AC AAY71169;

XX
DT 08-SEP-2000 (first entry)

XX
DE Adeno-associated virus serotype 1 capsid protein VP3.

XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; KW vaccine; transgene; VP3.

XX
OS Adeno associated virus serotype 1.

XX
PN WO200028061-A2.

XX
PD 18-MAY-2000.

XX
PF 02-NOV-1999; 99WO-US25694.

XX
PR 05-NOV-1998; 98US-0107114.

XX
PA (UYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Xiao W;

XX
DR WPI; 2000-376571/32.

DR
N-PSDB; AAD00772, AAD00779.

XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for

PT preparation of medicament for delivery of a transgene to a host -
XX
PS Claim 7; Page 99-101; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 cap protein VP3 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 534 AA;

Query Match 100.0%; Score 2906; DB 21; Length 534;
Best Local Similarity 100.0%; Pred. No. 9.1e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 60
Db 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 60

QY 61 ASTGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLNINNNGFRPKRLNFKLFNIQVK 120
Db 61 ASTGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLNINNNGFRPKRLNFKLFNIQVK 120

QY 121 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 180
Db 121 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 180

QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 240
Db 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 240

QY 241 LYILNRTQNQSGAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 241 LYILNRTQNQSGAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300

QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMGSGVMI FGKESAGASNTALDNVMITD 360
Db 301 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMGSGVMI FGKESAGASNTALDNVMITD 360

QY 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGPWAKI 420

QY 421 PHTDGHFHPSPLMGGFGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTQVSVE 480
Db 421 PHTDGHFHPSPLMGGFGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTQVSVE 480

QY 481 IEWELQENSKRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 481 IEWELQENSKRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534

RESULT 2
AAY711168
ID AAY711168 standard; Protein; 599 AA.
XX
AC AAY711168;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP2.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP2.
XX
OS Adeno associated virus serotype 1.
XX WO200028061-A2.
PN
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00778.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host -
XX
PS Claim 7; Page 93-95; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 cap protein VP2 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 599 AA;

Query Match 100.0%; Score 2906; DB 21; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 60
Db 66 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 125

QY 61 ASTGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLNINNNGFRPKRLNFKLFNIQVK 120
Db 126 ASTGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLNINNNGFRPKRLNFKLFNIQVK 185

QY 121 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 180
Db 186 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 245

QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 240
Db 246 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY 305

QY 241 LYILNRTQNQSGAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 306 LYILNRTQNQSGAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 365

QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMGSGVMI FGKESAGASNTALDNVMITD 360
Db 366 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMGSGVMI FGKESAGASNTALDNVMITD 425

QY 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 426 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGPWAKI 485

CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.

XX

SQ Sequence 736 AA;

Query Match 99.2%; Score 2884; DB 22; Length 736;
Best Local Similarity 99.1%; Pred. No. 1e-234;
Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 262

Qy 61 ASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK 322

Qy 121 EVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOQGYLTNL 180
Db 323 EVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOQGYLTNL 382

Qy 181 NGSQAVGRSSFFCYCLEYFPSQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMNLIDQY 240
Db 383 NGSQAVGRSSFFCYCLEYFPSQMLRTGNFTFSYTFEDVPHSSYAHQSGLDRLMNLIDQY 442

Qy 241 LYLLNRTQNQSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF 300
Db 443 LYLLNRTQNQSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF 502

Qy 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMMSGVMI FGKESAGASNTALDNVMITD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMMSGVMI FGKESAGASNTALDNVMITD 562

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 563 EEEIKATNPVATERFGTVAVNLQSSSTDPATGDVHMVWQDRDVLQGPWAKI 622

Qy 421 PHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPAEFSA TKFASFITQYSTGQVSVE 480
Db 623 PHTDGHFHPSPMLGGFGLKHPPPPQILIKNTVPANPAEFSA TKFASFITQYSTGQVSVE 682

Qy 481 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 683 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 736

RESULT 5
AAB59846

ID AAB59846 standard; Protein; 736 AA.

XX

AC AAB59846;

XX

DT 28-MAR-2001 (first entry)

XX

DE AAV3B capsid protein VP1.

XX

KW AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.

XX

OS Adeno associated virus.

XX

PN US6156303-A.

XX

PD 05-DEC-2000.

XX

PF 11-JUN-1997; 97US-0873168.

XX

PR 11-JUN-1997; 97US-0873168.

XX

PA (UNIW) UNIV WASHINGTON.

XX

PI Russell DW, Rutledge EA;
XX WPI; 2001-060164/07.
DR

XX

PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes -
XX Disclosure; Fig 2; 50pp; English.
XX

CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV3B).
CC AAV3B can be used to construct AAV viral vectors for use in gene therapy
CC for a range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.

XX

SQ Sequence 736 AA;

Query Match 87.1%; Score 2530; DB 22; Length 736;
Best Local Similarity 86.0%; Pred. No. 8.6e-205;
Matches 460; Conservative 27; Mismatches 46; Indels 2; Gaps 2;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNSSGNWHCHDSQWLGLDRVITTTSTRTWALPTYNHLYKQISS 262

Qy 61 ASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHFGYSTPWGYFDNRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK 321

Qy 121 EVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOQGYLTNL 180
Db 322 EVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMPVQGYLTNL 381

Qy 181 NGSQAVGRSSFFCYCLEYFPSQMLRTGNFTFSYTFEEVPHSSYAHQSGLDRLMNLIDQY 240
Db 382 NGSQAVGRSSFFCYCLEYFPSQMLRTGNNFQSYTFEDVPHSSYAHQSGLDRLMNLIDQY 441

Qy 241 LYLLNRTQ-NQSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF 299
Db 442 LYLLNRTQGTTS GTTNQSRLLFSQAGPQSMQLQARNWLPGPCYRQQRVSKTDANDNNNSNF 501

Qy 300 TWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMMSGVMI FGKESAGASNTALDNVMIT 359
Db 502 PWTAAASKYHLNGRDSLNVNPGPAMASHKODEKFFPMHGNLIFGKEGTTASNAELDNVMIT 561

Qy 360 DEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWAK 419
Db 562 DEEEIRTNPVATEQYGTVANLQSSNTAPTTRTVNDQGALPGMWQDRDVLQGPWAK 621

Qy 420 IPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPAEFSA TKFASFITQYSTGQVS 479
Db 622 IPHTDGHFHPSPMLGGFGLKHPPPPQIMIKNTVPANPPTTFSPAKFASFITQYSTGQVS 681

Qy 480 EIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 EIEWELQKENS KRWNPEIQYTSNYKSNVNVDFTVDTNGVSEPRPIGTRYLTRNL 736

RESULT 6
AAB59845

ID AAB59845 standard; Protein; 736 AA.

XX

AC AAB59845;

XX

DT 28-MAR-2001 (first entry)

XX

DE AAV3A capsid protein VP1.

XX

KW AAV3A; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; capsid protein VP1.
XX
OS Adeno associated virus.
XX
XX US6156303-A.
XX PN
XX PD
XX PD 05-DEC-2000.
XX
PF 11-JUN-1997; 97US-0873168.
XX
PR 11-JUN-1997; 97US-0873168.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Russell DW, Rutledge EA;
XX
XX WPI; 2001-060164/07.
DR
XX
PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes
XX
XX Disclosure; Fig 2; 50pp; English.
PS
XX
CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein VP1 of one such serotype (AAV3A).
CC AAV3A can be used to construct AAV viral vectors for use in gene therapy
CC for a range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell.
XX
SQ Sequence 736 AA;

Query Match 87.0%; Score 2528; DB 22; Length 736;
Best Local Similarity 85.8%; Pred. No. 1.3e-204;
Matches 459; Conservative 28; Mismatches 46; Indels 2; Gaps 2;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTSRTWALPTYNNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGLDRVITTSRTWALPTYNNHLYKQISS 262

Qy 61 ASTGASNDNHYFGYSTPWGYFDNFRFHCHFSPRDWQRLINNNGFRPKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHYFGYSTPWGYFDNFRFHCHFSPRDWQRLINNNGFRPKKLSFKLFNIQVR 321

Qy 121 EVTTNDGVTTIANNLTSTVQVESDYQLPYVLGSAHQGCLPFPADVFMPQYGYLTIN 180
Db 322 GVTQNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPFPADVFMPQYGYLTIN 381

Qy 181 NGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSYTFEEVPFHSSYAHSQSLDRMLNPLIDQY 240
Db 382 NGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSYTFEDVPFHSSYAHSQSLDRMLNPLIDQY 441

Qy 241 LYLLNRTQ-NQSGSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF 299
Db 442 LYLLNRTQGTSGTTNQSRLLFSPQAGPSMSLQARNWLPGPCYRQQRLSKTANDNNNSNF 501

Qy 300 TWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMIFGKESAGASNTALDNVMIT 359
Db 502 PWTAAASKYHLNGRDSLVPNGPPAMASHKDDKDFPMSGVNIIFGKEGTTASNAELDNVMIT 561

Qy 360 DEEBIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWAK 419
Db 562 DEEBIRTPNPVATEQYGTVANNLQSSNTAPTGTGVNHQALPGMWQDRDVLQGPWAK 621

Qy 420 IPHTDGHFHPSPLMGGFGLKNPPQIILKNTVPANPPAEFSATKFASFITQYSTGVSV 479
Db 622 IPHTDGHFHPSPLMGGFGLKHPPQIIMIKNTVPANPPPTTSPAKFASFITQYSTGVSV 681

Qy 480 EIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 682 EIEWELQKENSKRWNPEIQYTSNYKSNVNVDFTVDTNGVYSEPRPIGTRYLTRNL 736

RESULT 7

AAAG65794
ID AAG65794 standard; Protein; 533 AA.
XX
AC AAG65794;
XX
DT 11-FEB-2002 (first entry)
XX
DE Adeno-associated virus 2 (AAV-2) major coat protein VP3.
XX
KW Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytostatic;
KW inverted terminal repeat; nontropic; neuroprotective; antianemic; ITR;
KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV;
KW major coat protein; AAV-2; VP3.
XX
OS Adeno-associated virus 2.
XX
PN WO200168888-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07927.
XX
PR 14-MAR-2000; 2000US-189110P.
XX
PA (NEUR-) NEUROLOGIX INC.
XX
PI Xiao W, During MJ;
DR WPI; 2001-596912/67.
DR N-PSDB; AAI66974.
XX
PT Recombinant viral vector useful in improving gene therapy in a subject,
PT and for increasing efficiency of entry into a cell, comprises a
PT chimeric capsid having one non-native amino acid sequence and a desired
PT transgene
XX
PS Disclosure; Page 51; 53pp; English.
XX
CC The invention provides a recombinant viral vector (RVV) comprising a
CC chimeric capsid (I) having at least one non-native amino acid sequence,
CC derived from a capsid protein domain of parvovirus (II), a virus (III),
CC or their combination, and a transgene flanked 5' and 3' by inverted
CC terminal repeat (ITR) sequences, derived from (II), (III), or their
CC combination. The RVV is useful for improving gene therapy in a subject
CC with a disorder, and for increasing the efficiency of entry into a cell,
CC which involves producing (I) encapsulating a viral vector, and contacting
CC a cell with RVV having (I) such that (I) binds to an attachment site on
CC the cell surface and permits the vector to enter the cell efficiently.
CC A pharmaceutical composition comprising RVV with (I) containing a
CC transgene sequence associated with a disease or a disorder such that
CC expression of the transgene would result in amelioration of the disease
CC or disorder such as inherited neurological and metabolic diseases e.g.
CC lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polyneuropathy,
CC Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood
CC e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic
CC fibrosis, diabetes, diseases associated with hormone deficiencies,
CC retinoblastoma and various types of neoplastic cells which include
CC tumours especially central nervous system tumours, neoplasms, carcinomas,
CC sarcomas, leukemias and lymphoma. The present sequence represents the
CC adeno-associated virus 2 (AAV-2) major coat protein VP3. AAV-2 sequences
XX are used in the construction of a chimeric vector.
SQ Sequence 533 AA;

Query Match 85.4%; Score 2481.5; DB 22; Length 533;
Best Local Similarity 83.3%; Pred. No. 6.8e-201;

Matches	445;	Conservative	37;	Mismatches	51;	Indels	1;	Gaps	12
Qy	1	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISS	60						
Db	1	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDGRVITTTSTRTWALPTYNHLYKQISS	60						
Qy	61	ASTGASNDNHFGYSTPWGYFDENRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK	120						
Db	61	QS-GASNDNHFGYSTPWGYFDENRFHCHFSRDPDQRLINNNWGFPRKRLNFKLFNIQVK	119						
Qy	121	EVTITNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQGLPPPADVFEMIPOYGYLTNL	180						
Db	120	EVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGLPPPADVFEMVPOYGYLTNL	179						
Qy	181	NGSOAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVPPHSSVAHSQSGLDRLMNPIDQY	240						
Db	180	NGSOAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEDVPFPHSSVAHSQSGLDRLMNPIDQY	239						
Qy	241	LYYLNRNTONQSGSAQNDLLFSRGSPPAGMSVQPKWLPGPCYRQQRVSKTKTDNNNSNFT	300						
Db	240	LYYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS	299						
Qy	301	WTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMGVMIFGKESAGASNTALDVMITD	360						
Db	300	WTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVLIFGKQSEKTNVDIEKVMITD	359						
Qy	361	EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGVMWQDRDYYLQGPWAKI	420						
Db	360	EEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDYYLQGPWAKI	419						
Qy	421	PHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE	480						
Db	420	PHTDGHFHPSPLMGGFGLKHPPPPQILIKNTVPANPSTTFSAKFAFITQYSTGQVSVE	479						
Qy	481	IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEFRPICTRYLTRPL	534						
Db	480	IEWELQKENSKRWNPEIQYTSNYNKSANVDFTVDTNGVYSEPRPICTRYLTRNL	533						

RESULT 8
AAU98976
ID AAU98976 standard; Protein; 533 AA.
XX
AC AAU98976;
XX
DT 05-NOV-2002 (first entry)
XX
DE Adeno-associated virus 2 (AAV2) vector, VP3 capsid protein.
XX
KW Adeno-associated virus 2 vector; AAV2; cancer; VP3 capsid;
KW heparin-sulphate proteoglycan; vaccine; immune response;
KW ovarian cancer.
XX
OS Adeno-associated virus 2.
XX
PN WQ200253703-A2.
XX
PD 11-JUL-2002.
XX
PF 04-JAN-2002; 2002WO-US00152.
XX
PR 05-JAN-2001; 2001US-260124P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bartlett JS;
XX
DR WPI; 2002-583608/62.
DR N-PSDB; ABK89694.
XX
PT New adeno-associated virus vector comprises a biotinylated capsid or
PT capsid protein with an amino acid insertion in the VP1 capsid, useful
PT as a vaccine or for transferring a therapeutic peptide to a cancer cell.

-

Disclosure; Page 54-56; 57pp; English.

The invention relates to an adeno-associated virus (AAV) vector (I) comprising a biotinylated capsid or capsid protein (II) with an amino acid insertion following the capsid amino acid at position 139, 161, 588 or 657 in the VP1 capsid. The AAV vector comprises a capsid protein containing one or more amino acid insertions that ablate the ability of the vector to bind heparin-sulphate proteoglycan and allow the vector to use a cellular receptor not used by wild type AAV. Modified (I) are useful as vaccines to elicit immune responses to amino acids, where the response can be protective and/or therapeutic. (I) may be used to transfer a therapeutic peptide to a cancer cell, particularly to an ovarian cancer cell. The present sequence represents the adeno-associated virus 2 (AAV2) vector, VP3 capsid protein used in to make modified AAV2 vectors.

XX
SQ Sequence 533 AA;

Query Match 85.4%; Score 2481.5; DB 23; Length 533;
Best Local Similarity 83.3%; Pred. No. 6.8e-201;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGGAPMADNNEGADVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATGSAGPMADNNEGADVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 60

QY 61 ASTGASNHNHYFGYSTPWGYFDENRFCHSFPRDWQRLLNNNWGRPKRLNFKLFIQVK 120
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 QS-GASNHNHYFGYSTPWGYFDENRFCHSFPRDWQRLLNNNWGRPKRLNFKLFIQVK 119

QY 121 EVTTNDGVTTIANLNTSTVOVFSDSEYQLPYVLGSAHQGLPPFPADVFMIPOYGILTILN 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 EVTQNDGVTTIANLNTSTVOVFTDSEYQLPYVLGSAHQGLPPFPADVFMPVPOYGILTILN 179

QY 181 NGSQAVGRSSFYCLEYFPSQMLRTGNFTSYTFEEVPFHSSYAHQSQSLDRLMNPLIDQY 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 NGSQAVGRSSFYCLEYFPSQMLRTGNFTSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQY 239

QY 241 LYYLNRTNQSGSAQNKKOLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSFT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 LYYLSRTNPSTGTTTQSRLOFSQAGASDIRDSRNWLPGPCYRQQRVSKTSADNNNSEYS 299

QY 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNVMTD 360
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 WTGATKYHLNGRDSLNVNPGPAMASHKODEEKFFPSQGLVIFGKGSEKTNVDIEKVMTD 359

QY 361 EEEIKATNPVATERFGTVANVFQSSSTDPAATGDVHAMGALPGMWQDRDVYLQGPIWAKI 420
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 EEERIITNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKI 419

QY 421 PHTDGHFHPSPLMGGFGLKNPPPQILLIKNTPVPANPPAEFSATKFASFITOYSTGOVSVE 480
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 PHTDGHFHPSPLMGGFGLKHPPPQILLIKNTPVPANPFSTFSAKFASFITOYSTGOVSVE 479

QY 481 IEWELOKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 IEWELOKENSKRWNPEIOYTSNYNKSVNVDFTVDPTNGVYSEPRPIGTRYLTRNL 533

RESULT 9	
AAG65793	
ID	AAG65793 standard; Protein; 598 AA.
XX	
AC	
XX	AAG65793;
XX	
DT	11-FEB-2002 (first entry)
XX	
DE	Adeno-associated virus 2 (AAV-2) major coat protein VP2.
XX	
KW	Recombinant viral vector; RW; capsid; parvovirus; transgene; cytostatic;
KW	inverted terminal repeat; nootropic; neuroprotective; antianemic; ITR;

KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV;
KW major coat protein; AAV-2; VP2.
XX Adeno-associated virus 2.
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "encoded by ACG"
XX
XX WO200168888-A2.
PN 20-SEP-2001.
XX
XX 13-MAR-2001; 2001WO-US07927.
PF
XX
XX 14-MAR-2000; 2000US-189110P.
PR
XX (NEUR-) NEUROLOGIX INC.
PA
XX
PI Xiao W, During MJ;
XX
XX WPI; 2001-596912/67.
DR N-PSDB; AAI66974.
DR
XX
XX
PT Recombinant viral vector useful in improving gene therapy in a subject,
PT and for increasing efficiency of entry into a cell, comprises a
PT chimeric capsid having one non-native amino acid sequence and a desired
PT transgene -
XX
PS Disclosure; Page 51; 53pp; English.
XX
CC The invention provides a recombinant viral vector (RVV) comprising a
CC chimeric capsid (I) having at least one non-native amino acid sequence,
CC derived from a capsid protein domain of parvovirus (II), a virus (III),
CC or their combination, and a transgene flanked 5' and 3' by inverted
CC terminal repeat (Itr) sequences, derived from (II), (III), or their
CC combination. The RVV is useful for improving gene therapy in a subject
CC with a disorder, and for increasing the efficiency of entry into a cell,
CC which involves producing (I) encapsulating a viral vector, and contacting
CC a cell with RVV having (I) such that (I) binds to an attachment site on
CC the cell surface and permits the vector to enter the cell efficiently.
CC A pharmaceutical composition comprising RVV with (I) containing a
CC transgene sequence associated with a disease or a disorder such that
CC expression of the transgene would result in amelioration of the disease
CC or disorder such as inherited neurological and metabolic diseases e.g.
CC lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polynuropathy,
CC Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood
CC e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic
CC fibrosis, diabetes, diseases associated with hormone deficiencies,
CC retinoblastoma and various types of neoplastic cells which include
CC tumours especially central nervous system tumours, neoplasms, carcinomas,
CC sarcomas, leukemias and lymphoma. The present sequence represents the
CC adeno-associated virus 2 (AAV-2) major coat protein VP2. AAV-2 sequences
CC are used in the construction of a chimeric vector.
XX
SQ Sequence 598 AA;

Query Match 85.4%; Score 2481.5; DB 22; Length 598;
Best Local Similarity 83.3%; Pred. No. 8e-201;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDVRVITTTSTRTWALPTYNNHLYKQISS 60
Db 66 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDGRVITTTSTRTWALPTYNNHLYKQISS 125

Qy 61 ASTGASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRLINNWNWFRPKRLNFKLFNIQVK 120
Db 126 QS-GASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQRLINNWNWFRPKRLNFKLFNIQVK 184

Qy 121 EVTTNDGVTTIANNLTSTVQVFSDEXQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 180
Db 185 EVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTN 244

Qy 181 NGSQAVGRSSFYCLEYFPQSMLRTGNFTFSYTFEEVPFHSSYAHSQSLDRLMNLIDQY 240
Db 245 NGSQAVGRSSFYCLEYFPQSMLRTGNFTFSYTFEDVPFHSSYAHSQSLDRLMNLIDQY 304

Qy 241 LYYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSFT 300
Db 305 LYYLSRTNTPSGTTTOSRLQFSQAGASDIRDOSRNWLPGPCYRQQRVSKTSADNNNSEYS 364

Qy 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMGSGWMIIFGKESAGASNTALDNVMITD 360
Db 365 WTGATKYHLNGRDSLVPFGPAMASHKODEKFFPQSGVLIFGKQSGSEKTNVDIEKVMITD 424

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDQVYLOGPIWAKI 420
Db 425 EEEIIRTNPVATEQYGSVSTNLQRGNRQAATAADVNTQGVLPGMVWQDRDQVYLOGPIWAKI 484

Qy 421 PHTDGHFHPSPMLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTQGVSV 480
Db 485 PHTDGHFHPSPMLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKEFSITQYSTQGVSV 544

Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 545 IEWELQKENSKRWNPEIQYTSNYKSNVNDFTVDNNGVSEPRPIGTRYLTRNL 598

RESULT 10
AAM51509
ID AAM51509 standard; protein; 598 AA.
XX
AC AAM51509;
XX
DT 02-JAN-2002 (first entry)
XX
DE Adeno-associated virus VP2 capsid protein.
XX
KW Adeno-associated virus; AAV; VP2; capsid; virus-like particle;
KW nuclear localisation signal; VP3.
XX
OS Adeno associated virus.
XX
PN JP2001169777-A.
XX
PD 26-JUN-2001.
XX
PF 30-JUL-1999; 99JP-0249140.
XX
PR 30-JUL-1999; 99JP-0249140.
XX
PA (HAND/) HANDA H.
XX
DR WPI; 2001-599854/68.
XX
PT New virus-like particles from VP3 capsid protein of adeno-associated
PT virus, comprise a peptide containing a nucleus-shifting signal
PT connected to its N-terminal -
XX
PS Claim 1; Page 14-16; 33pp; Japanese.
XX
CC The present sequence is provided in a specification relating to a
CC virus-like particle-forming protein, and to a peptide containing a
CC nuclear-shifting signal at its N-terminus that can form virus-like
CC particles by shifting to the nucleus of the animal cell in which it
CC is expressed. The method is used for forming virus-like particles
CC from the VP3 capsid protein of adeno-associated virus (AAV).
XX
SQ Sequence 598 AA;

Query Match 85.4%; Score 2481.5; DB 22; Length 598;
Best Local Similarity 83.3%; Pred. No. 8e-201;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDVRVITTTSTRTWALPTYNNHLYKQISS 60

Db 66 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVTITSTRTWALPTVNNHLYKQISS 125

Qy 61 ASTGASNDNHYFGYSTPWGYFDNRFCHFSRDRWQRLNNWNGFRPKRLNFKLFNIQVK 120

Db 126 QS-GASNDNHYFGYSTPWGYFDNRFCHFSRDRWQRLNNWNGFRPKRLNFKLFNIQVK 184

Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMTPOYGYLTNL 180

Db 185 EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMTPOYGYLTNL 244

Qy 181 NGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPHSSYAHSSQSLDRLMNLIDQY 240

Db 245 NGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEDVPHSSYAHSSQSLDRLMNLIDQY 304

Qy 241 LYTLNRTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSNFT 300

Db 305 LYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNNSEYS 364

Qy 301 WTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGVMI FGKESAGASNTALDNVMTD 360

Db 365 WTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLI FGKQSEKTNVDIEKVMITD 424

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSDPATGDVHAMGALPGMWQDRDVLQGPWAKI 420

Db 425 EEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 484

Qy 421 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTQVSV 480

Db 485 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPSTTFSAKFASFITQYSTQVSV 544

Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534

Db 545 IEWELQKENSKRWNPEIQYTSNYKSNVNDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 11

AAU98975

ID AAU98975 standard; Protein; 598 AA.

XX AC AAU98975;

XX 05-NOV-2002 (first entry)

DE Adeno-associated virus 2 (AAV2) vector, VP2 capsid protein.

XX Adeno-associated virus 2 vector; AAV2; cancer; VP2 capsid;

KW heparin-sulphate proteoglycan; vaccine; immune response;

KW ovarian cancer.

XX Adeno-associated virus 2.

OS WO200253703-A2.

XX 11-JUL-2002.

PP 04-JAN-2002; 2002WO-US00152.

XX 05-JAN-2001; 2001US-260124P.

PR (CHIL-) CHILDRENS HOSPITAL INC.

XX Bartlett JS;

PI WPI; 2002-583608/62.

XX N-PSDB; ABK89694.

DR New adeno-associated virus vector comprises a biotinylated capsid or

XX capsid protein with an amino acid insertion in the VP1 capsid, useful

PT as a vaccine or for transferring a therapeutic peptide to a cancer cell

PT -

XX Disclosure; Page 52-54; 57pp; English.

PS

XX

CC The invention relates to an adeno-associated virus (AAV) vector (I) comprising a biotinylated capsid or capsid protein (II) with an amino acid insertion following the capsid amino acid at position 139, 161, 588 or 657 in the VP1 capsid. The AAV vector comprises a capsid protein containing one or more amino acid insertions that ablate the ability of the vector to bind heparin-sulphate proteoglycan and allow the vector to use a cellular receptor not used by wild type AAV. Modified (I) are useful as vaccines to elicit immune responses to amino acids, where the response can be protective and/or therapeutic. (I) may be used to transfer a therapeutic peptide to a cancer cell, particularly to an ovarian cancer cell. The present sequence represents the adeno-associated virus 2 (AAV2) vector, VP2 capsid protein used in to make modified AAV2 vectors.

XX SQ Sequence 598 AA;

Query Match 85.4%; Score 2481.5; DB 23; Length 598;
Best Local Similarity 83.3%; Pred. No. 8e-201;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITSTRTWALPTVNNHLYKQISS 60

Db 66 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVTITSTRTWALPTVNNHLYKQISS 125

Qy 61 ASTGASNDNHYFGYSTPWGYFDNRFCHFSRDRWQRLNNWNGFRPKRLNFKLFNIQVK 120

Db 126 QS-GASNDNHYFGYSTPWGYFDNRFCHFSRDRWQRLNNWNGFRPKRLNFKLFNIQVK 184

Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMTPOYGYLTNL 180

Db 185 EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMTPOYGYLTNL 244

Qy 181 NGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVPHSSYAHSSQSLDRLMNLIDQY 240

Db 245 NGSQAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEDVPHSSYAHSSQSLDRLMNLIDQY 304

Qy 241 LYTLNRTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSNFT 300

Db 305 LYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNNSEYS 364

Qy 301 WTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGVMI FGKESAGASNTALDNVMTD 360

Db 365 WTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLI FGKQSEKTNVDIEKVMITD 424

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSDPATGDVHAMGALPGMWQDRDVLQGPWAKI 420

Db 425 EEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 484

Qy 421 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTQVSV 480

Db 485 PHTDGHFHPSPMLMGFGGLKNPPQILIKNTVPANPSTTFSAKFASFITQYSTQVSV 544

Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534

Db 545 IEWELQKENSKRWNPEIQYTSNYKSNVNDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 12

AAG65792

ID AAG65792 standard; Protein; 735 AA.

XX AC AAG65792;

XX 11-FEB-2002 (first entry)

DE Adeno-associated virus 2 (AAV-2) major coat protein VP1.

XX Recombinant viral vector; RW; capsid; parvovirus; transgene; cytostatic;

KW inverted terminal repeat; neurotropic; neuroprotective; antianemic; ITR;

KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV;

KW major coat protein; AAV-2; VP1.

XX Adeno-associated virus 2.

Db 322 EVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCCLPPFPADVFMVPOYGYLTNL 381

Qy 181 NGSQAVGRSSFCLEYFPSQMLRTGNFTFSYTFEEVFPFHSSYAHQSLSLDRLMNPLIDQY 240
|||||

Db 382 NGSQAVGRSSFCLEYFPSQMLRTGNFTFSYTFEEVFPFHSSYAHQSLSLDRLMNPLIDQY 441
|||||

Qy 241 LYYLNRQTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
|||||

Db 442 LYYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 501
|||||

Qy 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNNVMTD 360
|||||

Db 502 WTGATKYHLNGRDSLVPNPGPAMASHKODEKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 561
|||||

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKI 420
|||||

Db 562 EEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 621
|||||

Qy 421 PHTDGHFHPSPMLMGFGGLKHPPPQILIKNTVPANPAEFSAATKFAFITQYSTGQVSVE 480
|||||

Db 622 PHTDGHFHPSPMLMGFGGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVE 681
|||||

Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
|||||

Db 682 IEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 735
|||||

RESULT 14
AAB59844

ID AAB59844 standard; Protein; 735 AA.

AC AAB59844;

XX 28-MAR-2001 (first entry)

DE AAV2 capsid protein VP1.

XX AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
atherosclerosis; sickle cell anaemia; thalassaemia;
blood clotting disorder; diabetes; capsid protein VP1.

OS Adeno associated virus.

XX US6156303-A.

PN 05-DEC-2000.

XX 11-JUN-1997; 97US-0873168.

XX 11-JUN-1997; 97US-0873168.

PR (UNIW) UNIV WASHINGTON.

XX Russell DW, Rutledge EA;

PI WPI; 2001-060164/07.

XX Adeno-associated virus serotype 6 and viral vector derived from it for
gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
syndrome, sickle cell anemia, thalassemia and diabetes -

PS Claim 7; Fig 2; 50pp; English.

XX The present invention relates to adeno-associated virus serotypes. The
present sequence is capsid protein VP1 of one such serotype (AAV2). AAV2
can be used to construct AAV viral vectors for use in gene therapy for a
range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
The AAV viral vectors have increased transduction efficiency of a
particular host cell as the AAV virion containing the AAV vector genome
can be modified to express a capsid protein of an AAV serotype that
transduces the selected host cell.

SQ Sequence 735 AA;

Query Match 85.4%; Score 2481.5; DB 22; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.1e-200;
Matches 445; Conservative 37; Mismatches 51; Indels 1;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVITSTRTWALPTYNNHLYKQISS 60
|||||

Db 203 MATGSGAPMADNNEGADGVGNSSGNWCHDSTWMGDRVITSTRTWALPTYNNHLYKQISS 262
|||||

Qy 61 ASTGASNDNHVFGYSTPWGYFDNRFHCHFSRDRWQRLINNNMGRPKRLNFKLFNIQVK 120
|||||

Db 263 QS-GASNDNHVFGYSTPWGYFDNRFHCHFSRDRWQRLINNNMGRPKRLNFKLFNIQVK 321
|||||

Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCCLPPFPADVFMVPOYGYLTNL 180
|||||

Db 322 EVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCCLPPFPADVFMVPOYGYLTNL 381
|||||

Qy 181 NGSQAVGRSSFCLEYFPSQMLRTGNFTFSYTFEEVFPFHSSYAHQSLSLDRLMNPLIDQY 240
|||||

Db 382 NGSQAVGRSSFCLEYFPSQMLRTGNFTFSYTFEDVPFHSSYAHQSLSLDRLMNPLIDQY 441
|||||

Qy 241 LYYLNRQTQNSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
|||||

Db 442 LYYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 501
|||||

Qy 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNNVMTD 360
|||||

Db 502 WTGATKYHLNGRDSLVPNPGPAMASHKODEKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 561
|||||

Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKI 420
|||||

Db 562 EEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 621
|||||

Qy 421 PHTDGHFHPSPMLMGFGGLKHPPPQILIKNTVPANPAEFSAATKFAFITQYSTGQVSVE 480
|||||

Db 622 PHTDGHFHPSPMLMGFGGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVE 681
|||||

Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
|||||

Db 682 IEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 735
|||||

RESULT 15
AAU98974

ID AAU98974 standard; Protein; 735 AA.

XX AAU98974;

XX 05-NOV-2002 (first entry)

DE Adeno-associated virus 2 (AAV2) vector, VP1 capsid protein.

XX Adeno-associated virus 2 vector; AAV2; cancer; VP1 capsid;
heparin-sulphate proteoglycan; vaccine; immune response;
ovarian cancer.

OS Adeno-associated virus 2.

XX WO200253703-A2.

PN 11-JUL-2002.

XX 04-JAN-2002; 2002WO-US00152.

XX 05-JAN-2001; 2001US-260124P.

XX (CHIL-) CHILDRENS HOSPITAL, INC.

PI Bartlett JS;

XX WPI; 2002-583608/62.

DR N-PSDB; ABK89694.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: January 21, 2004, 16:00:42 ; Search time 12.5714 Seconds
(without alignments)
4084.983 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEFRPIGTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	
1	2177.5	74.9	504	1 VCPV3A	coat protein - ade	
2	1676	57.7	732	2 S52210	coat protein VP1 -	
3	517	17.8	673	1 VCPVB5	coat protein VP1 -	
4	487.5	16.8	781	1 VCPV19	coat protein VP1 -	
5	234.5	8.1	723	1 VCPVPP	coat protein VP1 -	
6	232.5	8.0	729	1 VCPVNA	coat protein VP1 -	
7	232.5	8.0	729	1 A60006	coat protein VP1 -	
8	229	7.9	587	1 B44276	coat protein VP1 -	
9	226	7.8	722	1 VCPVV2	coat protein VP1 -	
10	197	6.8	716	1 VCPV2M	coat protein VP1 -	
11	196.5	6.8	718	1 VCPVIM	coat protein VP1 -	
12	194	6.7	722	1 VCPVME	coat protein VP1 -	
13	194	6.7	727	1 VCPV1F	coat protein VP1 -	
14	189	6.5	584	2 S49594	capsid protein VP2	
15	189	6.5	727	1 VCPVFP	coat protein VP1 -	
16	188	6.5	737	1 VCPVCD	coat protein VP1 -	
17	184	6.3	748	1 VCPVCP	coat protein VP1 -	
18	177.5	6.1	702	1 VCPVAP	coat protein VP1 -	
19	175	6.0	722	1 VCPVCN	coat protein VP1 -	
20	123.5	4.2	690	2 AB0124	probable TonB-depe	
21	122.5	4.2	880	1 SYBSVS	valine-tRNA ligase	
22	121	4.2	635	2 F96660	protein F2K11.10 [
23	118	4.1	739	2 T52289	probable transkeo	
24	115.5	4.0	345	1 G97024	probable phosphoes	
25	113.5	3.9	2817	2 B97033	uncharacterized pr	
26	109	3.8	418	2 T35753	probable periplasm	
27	109	3.8	1072	2 A86827	hypothetical prote	
28	109	3.8	1186	2 T18210	delta endotoxin -	
29	108.5	3.7	1711	2 AB1283	peptidoglycan link	

30 108 3.7 655 1 ALKBG cyclomaltodextrin
31 108 3.7 772 1 A44052 outer layer protei
32 108 3.7 826 2 AD1683 penicillin-binding
33 107.5 3.7 413 2 T23098 hypothetical prote
34 107.5 3.7 648 2 S50856 whn protein - rat
35 107 3.7 777 2 AE0462 probable exported
36 107 3.7 1338 2 T30565 MAP kinase kinase
37 106.5 3.7 403 2 B87513 acyl-CoA dehydroge
38 106 3.6 1271 2 D64237 hypothetical prote
39 105.5 3.6 2500 2 G71609 hypothetical prote
40 105 3.6 1394 2 A29637 position-specific
41 104.5 3.6 559 2 S33724 transcription fact
42 104.5 3.6 717 2 H85171 DEF (CLAL) protein
43 104 3.6 395 2 S50986 MAF1 protein - yea
44 104 3.6 715 2 H90977 hypothetical prote
45 104 3.6 719 2 F85824 hypothetical prote

ALIGNMENTS

RESULT 1
VCPV3A
coat protein - adeno-associated virus type 2
C;Species: adeno-associated virus type 2
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Jul-1999
C;Accession: A03698
R;Srivastava, A.; Lusby, E.W.; Berns, K.I.
J; Virol. 45, 555-564, 1983
A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.
A;Reference number: A03694; MUID:83164299; PMID:6300419
A;Accession: A03698
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-504 <SRI>
A;Cross-references: EMBL:J01901; NID:g209616; PIDN:AAA42376.1; PID:g209621; EMBL:M12405;
C;Superfamily: adeno-associated virus coat protein
C;Keywords: coat protein

Query Match		74.9%;	Score	2177.5;	DB	1;	Length	504;	
Best Local Similarity		82.5%;	Pred. No.	2.1e-153;					
Matches		392;	Conservative	33;	Mismatches	47;	Indels	3;	
						Gaps		2;	
QY	1	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGD	RVTITSTRTWALPTYNHLYKQISS	60					
Db	1	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMG	RVTITSTRTWALPTYNHLYKQISS	60					
QY	61	ASTGASNDNHYFGYSTPWGYFDENRFHCHFS	PRDWQRLINNNWGFRPKRLNFKLFNIQVK	120					
Db	61	QS-GASNDNHYFGYSTPWGYFDENRFHCHFS	PRDWQRLINNNWGFRPKRLNFKLFNIQVK	119					
QY	121	EVTNDGVTTIANLNTSTVQVFS	SEYQLPYVLGSAHQGCLPPFPADVFMI	PQYGYLTN	180				
Db	120	EVTQNDGTTTIANLNTSTVQVFTDSEYQLPYVLG	SAHQGCLPPFPADVFMPVQYGYLTN	179					
QY	181	NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSY	TFFEEVPHSSYAHQSGLDRLMNPLIDQY	240					
Db	180	NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSY	TFEDVPFHSSYAHQSGLDRLMNPLIDQY	239					
QY	241	LYYLNRTQNSGSAQNKDLLFSRGS	PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSFT	300					
Db	240	LYYLSRTNTPSGTTTQSRLOFSQAGASDIR	DQSRNWLPGPCYRQQRVSKTSADNNNSEYS	299					
QY	301	WTGASKYNLNGRESIINPGTAMASHKODE	KFFPMMSGVMI FGKESAGASNTALDNVMTD	360					
Db	300	WTGATKYHLNGRDSLVNP--AMASHKDD	EKFFPQSGVLI FGKQSEKTNVNIKVMITD	357					
QY	361	EEI KATNPVATERFGTVAVNFQSS	STD PATGDVHAMGALPGVMWQDRDVLQGP	IWAKI	420				
Db	358	EEI GTTNPVATEQYGSVSTNLQRGNR	QAATADVNTQGVLPGMWQDRDVLQGP	IWAKI	417				
QY	421	PHTDGHFHPSPMLGGFGLKNPPQILIK	NTVPANPPAEFSATKFPASFITQYSTG	475					

Db 418 PHTDGHFHPSPMLMGGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTG 472

RESULT 2

S52210

coat protein VP1 - muscovy duck parvovirus

N;Alternate names: VP1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000

C;Accession: S52210

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52210

A;Molecule type: DNA

A;Residues: 1-732 <ZAD>

A;Cross-references: EMBL:X75093; NID:g609091; PIDN:CAA52984.1; PID:g609093

A;Experimental source: strain FM

C;Genetics:

A;Gene: VP1

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

Query Match 57.7%; Score 1676; DB 2; Length 732;

Best Local Similarity 56.9%; Pred. No. 4.9e-116;

Matches 308; Conservative 77; Mismatches 142; Indels 14; Gaps 6;

QY 1 MASGGAPMADNNEGADVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISS 60

Db 199 MAEGGSGAMGDSAGGADGVGNASGNWHCDSQWLGLTVITKTTRTWLPSYNNHMYQAITS 258

QY 61 ASTGASNDNHFGYSTPWGYFDENRFHCHFSPRDWORLINNNWGRPKRLNPKLFNIQVK 120

Db 259 GTNPDSN-TOYAGYSTPWGYFDENRFHCHFSPRDWORLINNNWGRPKALPKFIFNVQVK 317

QY 121 EVTTNDGVTTIANNLTSTVQVFSQYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 180

Db 318 EVTTQDQTKTIANNLTSTQIFTDNEHQLPYVLGSAHQGCLPPFPADVFMIPQYGYCTMH 377

QY 181 ---NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEVPFHSSYAHSQSLDRLMPLI 237

Db 378 TNQSGARFNDRSFYCLEYFPSPQMLRTGNFTFSYTFEVPFHSMFAHSQSLDRLMPLI 437

QY 238 DQYLYLNRTONQSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTK--TDNN 295

Db 438 DQYLMNFESEV-NGGRNAQ-----FKKAVKGAFGAMGRNWLPGPKLLDQVRAYSGGTDNY 491

QY 296 NSNFTWTGASKYNLNGRESINPGTAMASHKDDDEDKFFPMSGVMI FGKE--SAGASNTAL 353

Db 492 ANWSIWSKGNKVFLKDREYLLQPGPVATTTHTEDQASSVPAQNIIGIAKDPYRSGSTLAGI 551

QY 354 DNVMITDEEEIKATNPVATERFGTVAWNFQSSSTDPAATGDVHAMGALPGMVWQDRDYLQ 413

Db 552 SDIMVTDEQEIAPTNGVGRPYGLTVTNEQNNTTAPTNAELXVLGALPGMVWQNRDIYLQ 611

QY 414 GPIWAKIPHTDGHFHPSPMLMGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYS 473

Db 612 GPIWAKIPKTDGKPHSPNLGGFGLHNPPPPQVFIKNTVPADPPLEYVNVQKNWSYITQYS 671

QY 474 TGQVSVIEIWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRP 533

Db 672 TGQCTVEMVWELRKENSKRWNPEIQFTSNFGNRTSTMFAPNETGGYVEDRLIGTRYLTQN 731

QY 534 L 534

Db 732 L 732

RESULT 3

VCPVB5

coat protein VP1 - bovine parvovirus

N;Contains: coat protein VP2

C;Species: bovine parvovirus

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A26104

R;Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.

J. Virol. 60, 1085-1097, 1986

A;Title: Complete nucleotide sequence and genome organization of bovine parvovirus.

A;Reference number: A26104; MUID:87061184; PMID:3783814

A;Accession: A26104

A;Molecule type: DNA

A;Residues: 1-673 <CHE>

A;Cross-references: EMBL:M14363; NID:g333454; PIDN:AAB59847.1; PID:g808805

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

F;138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 17.8%; Score 517; DB 1; Length 673;

Best Local Similarity 25.1%; Pred. No. 2.9e-30;

Matches 143; Conservative 85; Mismatches 208; Indels 134; Gaps 17;

QY 5 GGAPMADNNEGADVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISSASTG 64

Db 163 GSGSVGGGGRGGSGVGYSTGGTGTFISENIVVTNTRQFICDIKNGHLYKS-EVLNTG 221

QY 65 ASNDNHFGYSTPWGYFDENRFHCHFSPRDWORLINNNWGRPKRLNPKLFNIQKVEVTT 124

Db 222 DTAHRQY-AITTPWSYFNFNQYSSHSFSPDWQHLVNDYERFRPKAMIVRVYNLQIKQIMT 280

QY 125 NDGVTTIANN-LTSTVQVFSQYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLT----- 179

Db 281 DGAMGTVYNNDLTAGMHIFCDGHRYPYVQHPWDDQCMPENLSIWELPQYAYIPAPISV 340

QY 180 --NNGSQAVGRS-----SFYCLEYFPSPQMLRTGNFTFSYTFEVPFHSSYAHSQSLDRL 232

Db 341 VDNNTTNTVVEHLLKGVPLYMLENSDHEVLRNG-----RI 375

QY 233 MNPLIDQYLYLNRTON-----QSGSAQNKDLLFSRGS-----AGMSVQPKNWL 277

Db 376 YRIYIQLWRLMDRKQHHIQHASDDVQSTGQKQKLLIORTKQPNKQRFQNAALRTSNWM 435

QY 278 PGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMSG 337

Db 436 SGP-----GIARGTHNATLQTSAGALVTMTVNGAD-----VSG 469

QY 338 VM-----IFGKE-----SAGASNTALDNVMITDEEEIKATNPVAT 372

Db 470 VRAVRVGYSTDPIYGGQPPESDLLRLYSASAAEQQNPILEN-----AAR 515

QY 373 ERFGTVAWNFQSSSTDPAATGDVHAMGALPGMVWQDRDYLQGIWAKIPHTDGHFHPSP 432

Db 516 HTFTREATKLITSGNGADGDYKEWMLPNQMWDSAPISRYNPIWVKVPRVNRKTLTLD 575

QY 433 MGGFGLKNPPPPQILIK--NTPVPANPPAEFSATKFASFITQYSTGQVSVIEIWELOKENS 490

Db 576 DGSIPMSHPPTGTFIKLARIPVPGNGD-----SFLNIYVTGQVSCVWVEVERGRT 626

QY 491 KRWNPEVQYTSNYAKSANVD-FTVDNNGLY 519

Db 627 KWRPEYMH-----ATNMSVDAYTINNAGVY 653

RESULT 4

VCPV19

coat protein VP1 - parvovirus B19 (strain Au)

C;Species: parvovirus B19

A;Note: host Homo sapiens (man)

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C;Accession: A24299

R;Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.

J. Virol. 58, 921-936, 1986

A;Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated from

A;Reference number: A24299; MUID:86200451; PMID:3701931

A;Accession: A24299

A;Molecule type: DNA

A;Residues: 1-781 <SHA>


```
QY      3 SGGGAPMADNNEGADGVGNASG--NWHCDSTWLGD---RVITTSTRTWALPTYNNHLYKQ 57
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     174 SGGGGG-GGGRGAGGVGVSTGTFNNQTEFQYLGEGLVRITAHASRLIHLNMPHEHETYKR 232
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      58 IS--SASTGAS---NDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFPRKRLN 111
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     233 IHVLNSESAGVAGQVQDDAHTQMTVPWSLIDANAWGVWFNPADWQLISNNMTEINLVSFE 292
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     112 FKLFNIOQKEVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGCLPPFPADV 168
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     293 QEIFNVVLKTIITESATSPPTKIYNNDLTASLMVALDTNNTLPTTPAAPRSETLGFYPWLP 352
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     169 FMIPQYGYL-----TLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNMF 208
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     353 TKPTQYRYLSCIRNLNPPPTYTGQSQQITDSIQTGLHSDIMFYTIENAVPIHLLRTGDEF 412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     209 -TPSYTFEEVPHSSYAHQSQSLDR-----LMNPLI--DQYLYLNRTQNQSGSAQNK 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     413 STGIYHFDTKPL--KLTHSWQTNRSRLGLPPKLLTEPTTEGDQHPGTLPAANTRKGYHQT 470
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     258 DLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIIN 317
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     471 NNSYTEAT-----AIRP-----AQVGYNTPYMNFYSNGGPF-----LT 504
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     318 P--GTAMASHKDEDEKFFPMMSGVMIFGKESAGASNTALD--NVMITDEEEK--ATNP- 369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     505 PIVPTADTQYNDDEPN-----GAIRFTMDYQHGHLTSSQEELERYTFNPNQ 549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     370 -----VATERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDV 410
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     550 SKCGRAPKQOFNQOAPLNLENTNNGTLLPSDPIGKSNMHFMNTLNTYGPLTALNNTAPV 609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     411 YLQGPWAKIPHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPAEFSA-TKFA 466
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     610 FPNQGIWDKELDTDLKPRLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQQ 663
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     467 SFITQYSTGVSVIEWELQKENSQRNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 526
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     664 PRIITYSNFWWKGTLTFTAKWRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
A60006
coat protein VP1 - porcine parvovirus (strain 90HS)
N;Contains: coat protein vp2
C;Species: porcine parvovirus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C;Accession: A60006
R;Sakurai, M.; Nishimori, T.; Ushimi, C.; Nakajima, H.
Virus Res. 13, 79-86, 1989
A;Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.
A;Reference number: A60006; MUID:89319168; PMID:2750278
A;Accession: A60006
A;Molecule type: DNA
A;Residues: 1-729 <SAK>
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein; glycoprotein
F;151-729/Product: coat protein vp2 #status predicted <VP2>
F;172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #statu
```

```
Query Match      8.0%; Score 232.5; DB 1; Length 729;
Best Local Similarity 23.0%; Pred. No. 3.8e-09;
Matches 137; Conservative 86; Mismatches 245; Indels 127; Gaps 31;

QY      3 SGGGAPMADNNEGADGVGNASGNWH--CDSTWLGD---RVITTSTRTWALPTYNNHLYKQ 57
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     174 SGGGGG-GGGRGAGGVGVSTGTFNNQTEFQYLGEGLVRITAHASRLIHLNMPHEHETYKR 232
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      58 IS--SASTGAS---NDNHYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFPRKRLN 111
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     233 IHVLNSESAGVAGQVQDDAHTQMTVPWSLIDANAWGVWFNPADWQLISNNMTEINLVSFE 292
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     112 FKLFNIOQKEVT---TNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGCLPPFPADV 168
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db     293 QEIFNVVLKTIITESATSPPTKIYNNDLTASLMVALDTNNTLPTTPAAPRSETLGFYPWLP 352
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     169 FMIPQYGYL-----TLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNMF 208
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     353 TKPTQYRYLSCIRNLNPPPTYTGQSQQITDSIQTGLHSDIMFYTIENAVPIHLLRTGDEF 412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     209 -TPSYTFEEVPHSSYAHQSQSLDR-----LMNPLI--DQYLYLNRTQNQSGSAQNK 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     413 STGIYHFDTKPL--KLTHSWQTNRSRLGLPPKLLTEPTTEGDQHPGTLPAANTRKGYHQT 470
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     258 DLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIIN 317
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     471 NNSYTEAT-----AIRP-----AQVGYNTPYMNFYSNGGPF-----LT 504
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     318 P--GTAMASHKDEDEKFFPMMSGVMIFGKESAGASNTALD--NVMITDEEEK--ATNP- 369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     505 PIVPTADTQYNDDEPN-----GAIRFTMDYQHGHLTSSQEELERYTFNPNQ 554
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     370 VATERFGTVA-VNFQSSS-----TDPATG--DVHAMGALP-----GMVWQDRDVYLQGP 415
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     555 APKQOFNQOAPLNLENTNNGTLLPSDPIGKSNMHFMNTLNTYGPLTALNNTAPVFPNGQ 614
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     416 IWAKIPHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPANPPAEFSA-TKFA 471
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     615 IWDKELDTDLKPRLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFNADSPQQPRIIT 668
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     472 YSTGVSVIEWELQKENSQRNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 526
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     669 YSNFWWKGTLTFTAKWRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
B44276
coat protein VP1 - parvovirus LuIII
C;Species: parvovirus LuIII
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C;Accession: B44276
R;Diffoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A;Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A;Reference number: A44276; MUID:93297126; PMID:8517025
A;Accession: B44276
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-587 <DIF>
A;Cross-references: GB:M81888
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein; glycoprotein
F;49,90,220,304,371,503,511,514,539,571/Binding site: carbohydrate (Asn) (covalent) #statu
```

```
Query Match      7.9%; Score 229; DB 1; Length 587;
Best Local Similarity 22.7%; Pred. No. 5e-09;
Matches 133; Conservative 74; Mismatches 228; Indels 150; Gaps 28;

QY      3 SGGGAPMADNNEGADGVGNASGNWHCDSD--TWLGD---RVITTSTRTWAL--PTYNN--- 52
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     30 SGGG-----GSGGGGVGVSTGSDYDNQTHYKPLGDGWVEITAYSTRMVHLMMPKSENYCR 83
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     53 ---HLYKQISSASTGASNDNHYFGYSTPWGYFDNRFCHFSRPRDWQRLINNNWGFPRPKR 109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     84 VRVHNTNDTGTASHMAMDDEAHEQIW-TPWSLV DANAWGVWFQPSDWQYISNNMIHNLHS 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     110 LNFKLENIQKVEVT-TNDGVTTIA-----NNLTSTVQVFSSEYQLPYVLGSAHQGCL---- 161
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     143 LDQELFNVVVIKTVTEQNTGAEAIKVYNNDLTAAMVALDSNNILPYTPAIDNQETLGFYP 202
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     162 --PPFPA-----DVFMIPQYGYLTLNNGSQAVGRSSFYCYCLEYFPPS-QMLRTGN 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     203 WKPTIPSPYRYFSCDRNLSTVYKDEAGTITDTMGLASGLNSQFTTIENTQRINLLRTGD 262
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     207 NF-TFSYTFEEVPHSSYAHQSQSLDRMLNPLIDQ-----YLYLNRTQN- 249
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 263 EYATGYTFDEPIRLTHTTQWNRHLGQPPQITELPSSDTANATLTARGYRSLGTQIOGR 322

QY 250 -----QSGSAQNKD-LLFSRGSPAGMSVQPKNWLPG-----PCYRQOR 286

Db 323 NDVTEATVRPAQVGFQCPHDNFETSRAGPFKVPVVPADITQGLDHDANGSLRYTYDKQH 382

QY 287 VSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKFFPMGVMIFGKESA 346

Db 383 GQSWASQNNKDRYTW-DAVNYD-SGR----- 406

QY 347 GASNTALDNVMITDEEEIKATNPVATER---FGTVAVNFQS--SSTDPAATGDVHAMGALP 401

Db 407 WTNNCFIQSVVPTSEPN---ANQILTRDNLAGKTDIHFTNAFNSYGPLTAPPH----- 457

QY 402 GMVWQDRDVYLOGPIWAK---IPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTP--VPANP 456

Db 458 -----PARIYFQGIWDKELDLHKPRLHTQAPFV---CKNNAFGQLLVLRLAPNLTDQYD 509

QY 457 PAEFSATKFAFITQYSTGOVSVEIEWELQKENSKRWNPEVQYTS 501

Db 510 PNSSNLSRIVYGTFFWKGKLTLLKAK---MRPNA-TWNPVFQISA 550

RESULT 9

VCPVW2

coat protein VP1 - parvovirus H1

C;Species: parvovirus H1

A;Note: host Homo sapiens (man)

C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Apr-1994

C;Accession: A03699

R;Rhode III, S.L.; Paradiso, P.R.

J. Virol. 45, 173-184, 1983

A;Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization

A;Reference number: A03695; MUID:83112183; PMID:68233009

A;Accession: A03699

A;Molecule type: DNA

A;Residues: 1-722 <RHO>

A;Cross-references: EMBL:X01457; EMBL:J02198

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

Query Match 7.8%; Score 226; DB 1; Length 722;

Best Local Similarity 21.7%; Pred. No. 1.1e-08;

Matches 128; Conservative 83; Mismatches 234; Indels 146; Gaps 26;

QY 2 ASGGAPMADNNEGADGVGNASGNWHCDSTW--LGDRVITSTRTAL----- 47

Db 153 ADGGGS-SGGGGGGGGIGVSTGYDNTQTYKFLGDGWVEITAHASRLLLHLMPPSENYC 211

QY 48 --PTVNNHLYKQISSASTGASNDNHFGYSTPWPYGFDFNRHCHFSRDPWQRLINNMGF 105

Db 212 RVTVHNNQTTGHGTVKVGKNMAYDTHQQIW-TPWSLV DANAWGVWFQPSDWQFIQNSMESL 270

QY 106 RPKRLNFKLFNIQKEVTT-----NDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGC 160

Db 271 NLDSLQELFNVVVKTVTEQQGAGQDAIKVYNNDLTACVMVALDSNNILPYTPAAQTSET 330

QY 161 L-----PPFPAD---VFMIPOGYLTLNNGSQ-----AVG-----RSSFYCLE-YFPS 199

Db 331 LGFPYWKPTAPAPRYRYFFMPRQLSVTSSNSAEGTQITDTIGEPAQALNSQFFTIENLPI 390

QY 200 QMLRTGNNTF-SYTFEEVPHSSYA-----HSQSLDRLMN-PLIDQYLYLNRNQSG 252

Db 391 TLLRTGDEFTTGTYIFNTDPLKLTHTWQNRHLACLOGITDLPTSDTATASLTANGDRFG 450

QY 253 SAQNKDLLP-----SRGSPAGMSVQPKNWLPGPCYRQORVSK 289

Db 451 STQTQNVNYYTEALRTRPAQIGFMPQPHDNFEANRGGPFKVEVVP-----LDI 497

QY 290 TKTDNNNSNFTWTGASKYNLNGR--ESIINPGTA-----MASHKODEKFFPMMSG 337

Db 498 TAGEDHDAN----GAIRFNYGKHGEDWAKQGAAPERYTWDIDAIDSAAGRDTARCFV---- 549

QY 338 VMIFGKESAGASNTALDNVMITDEEEIKA-TNPVATERFGTVAVNFQSSSTDPAATGDVHA 396

Db 550 -----QSAPISIPPNQILQREDAIAGRTNMHYTNVNSYGPLSAFPHDP----- 596

QY 397 MGALPGMVWQDRDVYLOGPIWAK---IPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTPVP 453

Db 597 -----IYPNGQIWDKELDLHKPRLHTVAPFV---CKNPPGQLFVHLGP-- 638

QY 454 ANPPAEF--SATKFAFITQYSTGOVSVEIEWELQKENSKRWNPEVQYTSN 502

Db 639 -NLTDQFDPNSTTVSRIVT-YSTFYWKGIKFKAKLRPNLTWNPVYQATTD 687

RESULT 10

VCPV2M

coat protein VP1 - minute virus of mice

C;Species: minute virus of mice, murine parvovirus

C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Apr-1994

C;Accession: A03700

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03700

A;Molecule type: DNA

A;Residues: 1-716 <AST>

A;Cross-references: EMBL:V01115

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

Query Match 6.8%; Score 197; DB 1; Length 716;

Best Local Similarity 20.0%; Pred. No. 1.6e-06;

Matches 126; Conservative 88; Mismatches 241; Indels 174; Gaps 28;

QY 3 SGGAPMADNNEGADGVGNASGNWHCD--TWLGD---RVITTSTRTWTALPTVNNHLYKQ 57

Db 160 SGGG-----GSGGGGVGVSTGSYDNTQTHYRFLGDGWVEITATATRLVHLNMPKSENYCR 213

QY 58 I-----SSASTGASNDNHFGYSTPWPYGFDFNRHCHFSRDPWQRLINNMGFPKRL 110

Db 214 IRVHNTTDTSVKGNMAKDDAHEQIWTWPSLV DANAWGVWLQPSDWQYICNTMSQLNLVSL 273

QY 111 NFKLFNIQKEVTTND---GVTTIANNLSTVQVFSSEYQLPYVLGSAHQCLPPFPA 166

Db 274 DQEIFNVVLKTVTEQDLGGQAIKIYNNDLTACVMVALDSNNILPYTPAANSMETILGFYPW 333

QY 167 DVFMIPOGY-----LTLNNGSQAV-----GRSSFYCLEYFPS-QMLRTGNN 207

Db 334 KPTIASPYRYFCVDRDLSTVYENQEGTVEHNVMGTPKGIPOFFTIENTQITLRTGDE 393

QY 208 F-TFSYTFEEVPHSSYAHSQSLDRLMN--PLIDQYLYLNRNQSGSAQNKDLLFSRG 264

Db 394 FATGTYTFDTSV--KLHTWQNRQLGQPLLSTF-----PEADTDAGT-----LTAQG 441

QY 265 SPAGMSVQPKNWL-----PGCYRQORVSKTKTDNNNSNFT 300

Db 442 SRHGTQMGVNMVSEAIRTRPAQVGFQCPHDNFEASRAGP-FAAPKVPADITQGVDEAN 500

QY 301 WTGASKYNLNGRESIINPG-----TAMASHKODEKFF-----PMGVMIF 341

Db 501 GSVRYSYGKHGKNWASHGPAPERYTWDETSPGSRDTKDGFIQSAPLVVPPPLNGI--- 557

QY 342 GKSAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQS--SSTDPAATGDVHAMGA 399

Db 558 -----LTNANPIGTKN---DIHFSNVFNSYGLPTAFSH----- 587

QY 400 LPGMVWQDRDVYLOGPIWAK---IPHTDGHFHPSPLMGGFGLKNPPQILIKNTP--VPA 454

Db 588 -----PSPVYFQGIWDKELDLHKPRLHTVAPFV---CKNNAFGQLVRLGNLTDQ 637

QY 455 NPFAEFSATKFAFITQYSTGOVSVEIEWELQKENSKRWNPEVQYTSNYSANVDFTVD 514

Db 638 YDPNGATLSRIVYGTFFWKGKLTMRAKURA-----NTTWNPVYQVSAAE-----D 682

QY 515 NNGLY---TEPRIGT-----RYLTRPL 534
Db 683 NGNSYMSVTKWLPTATGNMQSVPLITRPV 711
RESULT 11
VCPVIM
coat protein VP1 - minute virus of mice (strain MVMi)
C;Species: minute virus of mice, murine parvovirus
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 28-Jul-2000
C;Accession: B23008; B29510
R;Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A;Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A;Reference number: A23008; MUID:85242059; PMID:3855242
A;Accession: B23008
A;Molecule type: DNA
A;Residues: 1-718 <SAH>
A;Cross-references: EMBL:X02481; NID:G60918; PIDN:CAB46507.1; PID:G5419928
R;Astell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A;Title: DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and
A;Reference number: A29510; MUID:86115415; PMID:3502703
A;Accession: B29510
A;Molecule type: DNA
A;Residues: 1-143, 'A', 145-718 <AST>
A;Cross-references: EMBL:M12032
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
Query Match 6.8%; Score 196.5; DB 1; Length 718;
Best Local Similarity 20.5%; Pred. No. 1.7e-06;
Matches 122; Conservative 82; Mismatches 243; Indels 147; Gaps 25;
QY 3 SGGGAPMADNNEGADVGNASGNWHCHDS--TWLGD---RVITTSTRTWALPTYNNHLYKQ 57
Db 161 SGGG-----GSGGGGVGVSTGSYDNQTHYRFLGDGWVEITALATRLVHLNMPKSENYCR 214
QY 58 I-----SSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPQRLINNMGFRPKRL 110
Db 215 IRVHNTTDTSVKGNMAKDDAHEQIWTWPSLVLDANAGVWLQPSDWQYICNTMSQLNLVSL 274
QY 111 NFKLFNIQKVEVTTND----GVTTIANNLSTVQVFSDEYQLPYVLGSAHQCLPPFPA 166
Db 275 DQEIFNVVLKTVTEQDSGGGAIKIYNNDLTACMVAVDSNNILPYTPAANSMETLGFYPW 334
QY 167 DVFMIPOYGY-----LTLNN-----GSAVGRSSFYCLEYFPS-QMLRTGN 206
Db 335 KPTIASPYRYFCVDRDLSVTYENQEGTIEHNVMGTPKGMNSQFFTIENTQITLLRTGD 394
QY 207 NP-TFSYTFEEVPHSSVAHSQSLDRLMN--PLIDQYLYLNRNTQNSGSAQNKDLLFSR 263
Db 395 EFATGTYFYFDTPV--KLTHWTQNRQLGQPPLLSTF-----PEADTDAGT-----LTAQ 442
QY 264 GSPAGMSVQPKNL-----PGPCYRQQRVSKTKTDNNNSNF 299
Db 443 GSRHGATQMEVNWVSEAIRTPAQVGFQCPHNDFEASRAGP-FAAPKVPADVTQGVDFEA 501
QY 300 TWTGASKYNLNGRESIINPGTAMASHKODEDKF-----FPMMSGVMI FGKESAGASN 350
Db 502 NGSVRYSYKQKHGENWAAHGPAPERYTWDETNFGSGRDRDGFQISAPLVVPPPLNGILT 561
QY 351 TALDNVMITDEEEIKATNPVATERFGTVAVNFQS--SSTDPAATGDVHAMGALPGMWQDR 408
Db 562 NA-----NPICTKN---DIHFSNVFNSYGPLTAFSH-----PS 591
QY 409 DVYLQGPWAK---IPHTDGHFHPSPMLGGFGLKNPPQIILIKNTP--VPANPPAEFSAT 463
Db 592 PVYPQGIWDKELDLHKPRLHITAPFV---CKKNAPGQMLVRLGPNLTDQYDENGATLS 648
QY 464 KPAFITYQYSTQVSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTTVDNNG 517

Db 649 RIVTYGTFFWKGKLTMRACLRA-----NTTWNPVYQ-----VSVEDNG 686
RESULT 12
VCPVME
coat protein VP1 - mink enteritis virus (strain Abashiri)
N;Contains: coat protein VP2
C;Species: mink enteritis virus, MEV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Apr-1996
C;Accession: B38350
R;Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinaga
J. Gen. Virol. 72, 867-875, 1991
A;Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: B38350
A;Molecule type: DNA
A;Residues: 1-722 <KAR>
A;Cross-references: GB:D00765
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;139-722/Product: coat protein VP2 #status predicted <VP2>
Query Match 6.7%; Score 194; DB 1; Length 722;
Best Local Similarity 20.7%; Pred. No. 2.7e-06;
Matches 136; Conservative 73; Mismatches 215; Indels 232; Gaps 32;
QY 5 GGAPMADNNEGADVGNASGNWHCHDSTWLGDRVITTS-----RTWALPTYNN- 52
Db 148 GGQP-AVRNERATSGNGSGGGGGGGS--GGVGISTGTFFNQTFFKFLENGWVEITANSS 204
QY 53 ---HL-----YKQI-----SSASTGASNDNHFGYSTPWGYFDNRFHCHFSRDPW 95
Db 205 RLVLHNMPESENYKRVVNVNMDKTAVKGNMALDDTHVQIVTPWSLVLDANAGVWFNPGDW 264
QY 96 ORLINNMGFRPKRLNFKLFNIQKVEVT--TNDGVTTIANNLSTVQVFSDEYQLPYV 152
Db 265 OLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPPT 324
QY 153 LGSAHQGL-----PPFPA-----DVFMIPOYGYLTLNNGSQAVGRS----- 189
Db 325 PAAMRSETLGFYPWKPTIPTPRVYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDD 378
QY 190 -SFYCLE-YFPSQMLRTGNF---TF-----SYTFEE-----VPFHSSYAHSQSL 229
Db 379 VQFYTIENSVPVHLLRTGDEFATGTFDFDCKPCRLTHTWQTNRAIGLPPPLNSLPQSEGA 438
QY 230 DR-----LMNPLIDQYL--YVLNRTONQS----- 251
Db 439 TNFGDIGVOQDKRRGVQTMGNTDYITEATIMRPAEVGYSAPYISFEASTQGFPTPIAAG 498
QY 252 -GSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCY---RQQRVSKTKTDNNNSNFTWTGASKY 307
Db 499 RGGAOQTDENQAADGDPR-----YAFGRHQHGQKTTTGTETPERFTY----- 538
QY 308 NLNGRESIINPGTAMASHKODEDKFPFMSGVMI FGKESAGASNTALD-NVMITDEEEIKA 366
Db 539 -----IAHQDT-----GRYPAGDWIQININFNLPVTNDNVLLP 570
QY 367 TNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKIPHTDGH 426
Db 571 TDPIG---GKTGINY--TNIFNTYGPLTALNNVP-----PVYPNGQIWDKBFDTD-- 615
QY 427 FHPSPMLGGFGLK-----NPPPIQLIKNTPVPAN---PPAEFSATKFAFIT 470
Db 616 -----LKPRLHVNAFPVCQNNCPGQLFVKVAPNLITNEYDPPASANMSR-----IV 660
QY 471 QYSTQVSVSEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTTVDNNGLYTEPRPIG 526
Db 661 TYSDFWWKGLVFKAKLRASHTWNPIQQMSIN-----VDNQFNYL-PNNIG 705
RESULT 13
VCPVLF

coat protein VP1 - feline panleukopenia virus

N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03701
R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A;Reference number: A03697; MUID:85265017; PMID:2991581
A;Accession: A03701

A;Molecule type: DNA
A;Residues: 1-727 <CAR>
A;Cross-references: EMBL:M10824; NID:G3333474; PIDN:AAA47161.1; PID:G3333476
C;Genetics: 11/1
A;Introns:
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 6.7%; Score 194; DB 1; Length 727;
Best Local Similarity 20.7%; Pred. No. 2.7e-06;
Matches 136; Conservative 73; Mismatches 215; Indels 232; Gaps 32;

QY	5	GGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTS	-----RTWALPTYNN- 52
Db	153	GGQP-AVRNERATSGNGSGGGGGGS--GGVGISTGTENNQTETKFLNGWVEITANSS	209
QY	53	---HL-----YKQI-----SSASTGASNDNHVFGYSTPWGYFDNRFHCHFSPRDW	95
Db	210	RLVHLNMPSENYKRVVNVNMDKTAVKGNMALDDTHVQIVTPWSLVDANAWGVWFNPGDW	269
QY	96	QRLNNWGFPRKRLNFKLNIQKVEV---TNDGVTTIANNLSTVQVFSSEYQLPYV	152
Db	270	QLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT	329
QY	153	LGSAHQGL-----PPFPA-----DVFMIPQYGYLTLLNNGSQAVGRS-----	189
Db	330	PAAMRSETLGFYPWKPTIPTPWRYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDD	383
QY	190	-SFYCLE-YFPQMLRTGNF---TF-----SYTFEE-----VPEHSSYAHSQSL	229
Db	384	VQFYTIENSVPVHLRLTGDEFATGTFDFCKPCRLTHTWQTNRALGLPPEFLNSLPQSEGA	443
QY	230	DR-----LMNPLIDQYL--YYLNRTQNS-----	251
Db	444	TNFGDIGVQDQKRGVTQMNTDYITEATIMRPAEVGSAPYYSFEASTQGFPIPIAAG	503
QY	252	-GSAQNKDLLPSRGSPAGMSVQPKNWLPGPCY---RQORVSKTKTDNNNSNFTWTGASKY	307
Db	504	RGGAQTDEQAADGPR-----YAFGRHQGQKTTTGTETPERFTY-----	543
QY	308	NLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFGKESAGASNTALD-NVMITDEEIIKA	366
Db	544	-----IAHQDT-----GRYPAGDWIQNINFNLPVTNDNVLLP	575
QY	367	TNPVATERFGTVAVNFQSSSTDPTATGDVHAMGALPGMVWQDRDYLQGPWAKIPHTDGH	426
Db	576	TDPIG-----GKTGINY--TNIFNTYGPLTALNNVP-----PVYPNGQIWDKEFTD--	620
QY	427	FHPSPLMGGFGLK-----NPPQILIKNTVPVAN---PPAEFSATKFASFIT	470
Db	621	-----LKPRLHVNAFPVCQNNCPQLFVKVAPNLTNEDPDASANMSR-----IV	665
QY	471	QYSTGQSVSEIWELOKENSKRWNPEVQYTSNYAKSANVDFTTVDNNGLYTEPRPIG	526
Db	666	TYSDFWKGKLVFKAKLRASHTWNPIQMSIN-----VDNQFNYL-PNNIG	710

RESULT 14
S49594
capsid protein VP2 - canine parvovirus
C;Species: canine parvovirus, CPV
C;Date: 05-Mar-1995 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999

C;Accession: S49594
R;Mizak, B.; Plucienniczak, A.
submitted to the EMBL Data Library, November 1994
A;Description: Polish isolates of canine parvovirus.
A;Reference number: S49594
A;Accession: S49594
A;Molecule type: DNA
A;Residues: 1-584 <MIZ>
A;Cross-references: EMBL:Z46651; NID:G572660; PIDN:CAA86612.1; PID:G572661
C;Superfamily: parvovirus coat protein

Query Match 6.5%; Score 189; DB 2; Length 584;
Best Local Similarity 20.8%; Pred. No. 4.6e-06;
Matches 132; Conservative 89; Mismatches 223; Indels 192; Gaps 33;

QY	5	GGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTS	-----RTWALPTYNN- 52
Db	10	GGQP-AVRNERATSGNGSGGGGGGS--GGVGISTGTENNQTETKFLNGWVEITANSS	66
QY	53	---HL-----YKQI-----SSASTGASNDNHVFGYSTPWGYFDNRFHCHFSPRDW	95
Db	67	RLVHLNMPSENYRRVVVNNLDKTAVNGNMAALDDTHAQIVTPWSLVDANAWGVWFNPGDW	126
QY	96	QRLNNWGFPRKRLNFKLNIQKVEV---TNDGVTTIANNLSTVQVFSSEYQLPYV	152
Db	127	QLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT	186
QY	153	LGSAHQGL-----PPFPA-----DVFMIPQYGYLTLLNNGSQAVGRS-----	189
Db	187	PAAMRSETLGFYPWKPTIPTPWRYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDD	240
QY	190	-SFYCLE-YFPQMLRTGNF-TFSYTFEEVFPFHSSVAHSQSLDRLMNLIDQYLYLNR	246
Db	241	VQFYTIENSVPVHLRLTGDEFATGTFDFCKP--CRLTHTWQTNRALG--LPPFLNSLPQ	296
QY	247	TQNS-----GSAQNK-----DLLFSRGSPAGMSVQPKNWLPGPCYRQQRV	287
Db	297	SEGTFNGYIGVQDQKRGVTQMNTNYITEATIMRPAEVGS-----APYYSFEAS	348
QY	288	SK-----TKTDNNNSNFTWTGASKY---NLNGRESIINPGT---AMASHKD	327
Db	349	TQGFPTPIAAGRGGAQTDEQA--DGDPRYAFGRHQGQKTTTGTETPERFTYIAHQD	405
QY	328	DEKFFPMGSMVIFGKESAGASNTALD-NVMITDEEIKATNPVATERFGTVAVNFQSSS	386
Db	406	T-----GRYPEGDWIQNINFNLPVTDNVLPTDPIG---GKTGINY--TN	446
QY	387	TDPATGDVHAMGALPGMVWQDRDYLQGPWAKIPHTDGHFHPSPLMGGFGLK-----	439
Db	447	IFNTYGPLTALNNVP-----PVYPNGQIWDKEFTD-----LKPRLHVNA	486
QY	440	-----NPPQILIKNTVPVAN---PPAEFSATKFASFITQYSTGQSVSEIWELOKENS	490
Db	487	PFVCQNNCPGQLFVKVAPNLTNEDPDASANMSR-----IVTYSDFWKGKLVFKAKLRAS	542
QY	491	KRWNPVQYTSNYAKSANVDFTTVDNNGLYTEPRPIG	526
Db	543	HTWNPIQMSIN-----VDNQFNYP-PSNIG	567

RESULT 15
VCPVFP
coat protein VP1 - feline panleukopenia virus (strain 193)
N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 12-Apr-1996
C;Accession: B36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvo
A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: B36608
A;Molecule type: DNA

Search completed: January 21, 2004, 16:07:09
Job time : 14.5714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 15:59:46 ; Search time 9.14286 Seconds
(without alignments)
2746.653 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177.5	74.9	504	1 COA3 AAV2	P03135 adeno-assoc
2	517	17.8	673	1 COAT_PAVBO	P07297 bovine parv
3	487.5	16.8	781	1 COAT_PAVHB	P07299 human parvo
4	232.5	8.0	729	1 COA1_PAVP9	P33484 porcine par
5	232.5	8.0	729	1 COA1_PAVPN	P18546 porcine par
6	230.5	7.9	729	1 COA1_PAVPK	P52501 porcine par
7	229	7.9	587	1 COAT_PAVL3	P36310 parvovirus
8	226	7.8	722	1 COAT_PAVHH	P03136 hamster par
9	197	6.8	716	1 COAT_MUMIV	P03137 murine minu
10	196.5	6.8	718	1 COAT_MUMIM	P07302 murine minu
11	194	6.7	722	1 COAT_MEVA	P27437 mink enteri
12	194	6.7	727	1 COAT_FPV	P04864 feline panl
13	190	6.5	584	1 COAT_PAVC2	P30129 canine parv
14	189	6.5	727	1 COAT_FPV19	P24840 feline panl
15	188	6.5	727	1 COAT_PAVCB	Q11213 canine parv
16	188	6.5	737	1 COAT_PAVCD	P17455 canine parv
17	184	6.3	748	1 COAT_PAVCN	P12930 canine parv
18	177.5	6.1	647	1 COAT_ADVG	P24029 aleutian mi
19	176	6.1	722	1 COAT_PAVC7	P04863 canine parv
20	122.5	4.2	648	1 FXN1_HUMAN	O15353 homo sapien
21	121	4.2	880	1 SYV_BACST	P11931 bacillus st
22	109	3.8	1186	1 CEAA_BACTS	Q45710 bacillus th
23	108	3.7	655	1 CDGT_KLEPN	P08704 klebsiella
24	107.5	3.7	648	1 FXN1_MOUSE	Q61575 mus musculu
25	106.5	3.7	1070	1 Y355_HUMAN	O15063 homo sapien
26	106	3.6	1271	1 Y338_MYCGE	P47580 mycoplasma
27	105	3.6	1396	1 ITA2_DROME	P12080 drosophila
28	104.5	3.6	559	1 HNF8_PIG	Q03365 sus scrofa
29	104.5	3.6	717	1 DXS_ARATH	Q38854 arabidopsis
30	104	3.6	395	1 MAE1_YEAST	P41910 saccharomyc
31	103.5	3.6	1196	1 AMYE_PAEPO	P21543 paenibacill
32	102	3.5	1122	1 ADP1_MYCGA	Q49379 mycoplasma
33	102	3.5	2493	1 CYAA_USTMA	P49606 ubtilago ma

34	102	3.5	3712	1 LMA_DROME	Q00174 drosophila
35	100.5	3.5	531	1 HEXB_PIG	Q29548 sus scrofa
36	100	3.4	739	1 CATA_MYCSM	Q59557 mycobacteri
37	99	3.4	879	1 GUN1_CLOTM	Q02934 clostridium
38	98.5	3.4	557	1 HNF8_RAT	P23899 rattus norv
39	97.5	3.4	560	1 GAT2_YEAST	P40209 saccharomyc
40	97.5	3.4	676	1 VGP_EBOIC	Q66810 ebola virus
41	97.5	3.4	685	1 RNF6_HUMAN	Q9Y252 homo sapien
42	97.5	3.4	1113	1 N116_YEAST	Q02630 saccharomyc
43	97.5	3.4	2233	1 COAC_YEAST	Q00955 saccharomyc
44	97	3.3	557	1 CNE6_HUMAN	O95741 homo sapien
45	97	3.3	800	1 GUN_BACSI	P06564 bacillus sp

ALIGNMENTS

RESULT 1

COA3_AAV2
ID COA3_AAV2 STANDARD; PRT; 504 AA.
AC P03135;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Probable coat protein 3.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83164299; PubMed=6300419;
RA Srivastava A., Lusby E.W., Berns K.I.;
RT "Nucleotide sequence and organization of the adeno-associated virus 2 genome.";
RL J. Virol. 45:555-564(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; J01901; AAA42376.1; --
DR PIR; A03698; VCPV3A.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
SQ SEQUENCE 504 AA; 56366 MW; 758999B017052B682 CRC64;

Query Match 74.9%; Score 2177.5; DB 1; Length 504;
Best Local Similarity 82.5%; Pred. No. 2.4e-149;
Matches 392; Conservative 33; Mismatches 47; Indels 3; Gaps 2;

QY	1	MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVTTSTRTWALPTYNHLYKQISS	60
DB	1	MATGSGAPMADNNEGADGVGNSSGNWCHDSTWMDRVTTTSTRTWALPTYNHLYKQISS	60
QY	61	ASTGASNDNHVFCYSTPWGYFDNFNFCHFSPRDWORLNNNWGFRPKLNFKNIOVK	120
DB	61	QS-GASNDNHVFCYSTPWGYFDNFNFCHFSPRDWORLNNNWGFRPKLNFKNIOVK	119
QY	121	EVTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMPQYGYLTN	180
DB	120	EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMPQYGYLTN	179
QY	181	NGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEEVPFHSSYAHSQSLDRLMNPIDQY	240
DB	180	NGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY	239
QY	241	LYLNRNTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT	300

Db 240 LYLSRNTPTSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 299

Qy 301 WTCASKYNLNGRESIINPGTAMASHKDDKEDKFFPMGVMIFGKESAGASNTALDNVMTD 360

Db 300 WTGATKYHLNGRDSLVP--AMASHKODEEKFFPQSGVLIIFGKGSEKTNVNIKVMITD 357

Qy 361 BEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPWAKI 420

Db 358 BEEIGTNPVATEQYGSVSTNLQRNQAATADVNTQGLVPGMVWQDRDVLQGPWAKI 417

Qy 421 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFAFITQYSTG 475

Db 418 PHTDGHFHPSPMLGGFGLKHPPQILIKNTVPANPSTTFSAKFASFITQYSTG 472

RESULT 2

COAT_PAVBO

ID COAT_PAVBO STANDARD; PRT; 673 AA.

AC P07297; Q84374;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Coat protein VP2 [Contains: Coat protein VP3].

OS Bovine parvovirus (BPV).

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OX NCBI_TaxID=10784;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87061184; PubMed=3783814;

RA Chen K.C., Shull B.C., Moses E.A., Lederman M., Stout E.R.,

RA Bates R.C.;

RT "Complete nucleotide sequence and genome organization of bovine parvovirus";

RL J. Virol. 60:1085-1097(1986).

CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; M14363; AAB59848.1; -.

DR EMBL; M14363; AAB59849.1; -.

DR PIR; A26104; VCPVB5.

DR HSSP; P30129; 4DPV.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

KW Coat protein; Glycoprotein.

FT CHAIN 1 673 COAT PROTEIN VP2.

FT CHAIN 138 673 COAT PROTEIN VP3.

FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DOMAIN 163 187 GLY-RICH.

SQ SEQUENCE 673 AA; 75103 MW; 5F244642B2214831 CRC64;

Query Match 17.8%; Score 517; DB 1; Length 673;

Best Local Similarity 25.1%; Pred. No. 1.1e-29;

Matches 143; Conservative 85; Mismatches 208; Indels 134; Gaps 17;

Qy 5 GGAPMADNNEGADGVGNAGNHCDSITWGLDRVITTTSTRTWALPTYNNHLYKQISSASTG 64

Db 163 GSGSVCGGGRGGSGVYSTGWTGCTIFSENIVVTNKTQFICDIKNGHLYKS-EVLNTG 221

Qy 65 ASNDNHFGYSTPWGYFDNRFHCHFSRDPWQRLINNNGWFRPKRLNFKLFNIQKEVTT 124

Db 222 DTAHRQY-AITTPWSYFNQYSSHSFSPNDWQHLVNDYFRFRPKAMIVRVNLIQKIMT 280

Qy 125 NDGVTTIANN-LTSTVQVFSDBEYQLPYVLGSAHQGLCPFPFADVFMIPOGYLTL---- 179

Db 281 DGAMGTVYNNDLTAGMHIFCDGDRYPYVQHPWDDQCMPELNPISIWELPQYAYIPAPISV 340

Qy 180 --NNGSQAVGRS-----SFYCLEYFPFQMLRTGNNFTFSYTFEEVFFHSSYAHQSQSLDRL 232

Db 341 VDNNTTNTVEEHLKGVPLYMLENSDHEVLNRG-----RI 375

Qy 233 MNPLIDQVLYLNRNTQN-----QSGSAQNKOLLFSRGP-----AGMSVQPKNWL 277

Db 376 YRIYIQLWRLMDRKQHHIQHASDDVQSTGQKQKLLIORTKQPNKQRFQNAALRTSNWM 435

Qy 278 PGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKEDKFFPMSG 337

Db 436 SGP-----GIARGTHNATLQTSAGALVTMTNGAD-----VSG 469

Qy 338 VM-----IFGKE-----SAGASNTALDNVMTDEEEIKATNPVAT 372

Db 470 VRAVRVGYSTDPYGGQOPESDLRLRYASAAEQONPILE-----AAR 515

Qy 373 ERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPWAKIPHTDGHFHPSP 432

Db 516 HTFTREARTKLITGSGNGADGDYKEWMLPNQMWDSAPISRYNPVWVPRVNRKTLTDTQ 575

Qy 433 MGGFGLKNPPQILIK--NTPVPANPPAEFSATKFAFITQYSTGQSVSEIEMELQKENS 490

Db 576 DGSIPMSHPGCTIFIKLARIPVPGNGD-----SFLNIYVTGQSVCEVWVEVEKRG 626

Qy 491 KRWNPVQVTSNYAKSANVD-FTVDNNGLY 519

Db 627 KWRPPEYMH---ATNMSVDAYTINNAGVY 653

RESULT 3

COAT_PAVHB

ID COAT_PAVHB STANDARD; PRT; 781 AA.

AC P07299;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-APR-1990 (Rel. 14, Last annotation update)

DE Probable coat protein VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Isolate AU;

RX MEDLINE=86200451; PubMed=3701931;

RA Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;

RT "Nucleotide sequence and genome organization of human parvovirus B19 isolated from the serum of a child during aplastic crisis.";

RL J. Virol. 58:921-936(1986).

CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; M13178; AAA66867.1; -.

DR PIR; A24299; VCPV19.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 781 AA; 86015 MW; 8C6254BDD0576B07 CRC64;

Query Match 16.8%; Score 487.5; DB 1; Length 781;

```
RESULT 5
COAL_PAVPN      ID COAL_PAVPN    STANDARD; PRT;   729 AA.
AC P18546; P22854; Q89816;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Porcine parvovirus (strain NADL-2) (PPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10797;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085785; PubMed=2596019;
RA Vasudevacharya J., Basak S., Srinivas R.V., Compans R.W.;
RT "Nucleotide sequence analysis of the capsid genes and the right-hand
terminal palindrome of porcine parvovirus, strain NADL-2.";
RL Virology 173:368-377(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021005; PubMed=2219713;
```



```
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 729 AA; 80835 MW; B6345BFA0568A1F6 CRC64;

Query Match 7.9%; Score 230.5; DB 1; Length 729;
Best Local Similarity 23.0%; Pred. No. 5.2e-09;
Matches 137; Conservative 86; Mismatches 245; Indels 127; Gaps 31;

QY 3 SGGGAPMADNNEGADGVGNASGNWH--CDSTWLGD--RVITSTRTWALPTYNHLYKQ 57
Db 174 SGGGGG-GGGGRGAGGVGVSTGSFNNQTEFQYLGEGLVRITAHASRLIHLNMPHEHYKR 232

QY 58 IS--SASTGAS----NDNHFGYSTPWGYFDNRFCHFSRDMQRLINNNWGRPKRLN 111
Db 233 IHLNSESAGVAGQVQDDAHTQMTWPSLIDANAWGVWENPADQLISNNMTEINLVSE 292

QY 112 FKLFIQVKEVT---TNDGVTTIANLITSTQVFSSEYQLPYVLGSAHQGLPPFPADV 168
Db 293 QEIFNVVLKXITESATSPPTKIYNNDLTASLMVALDTNNTLPTPAAPRSETLGFYFWLP 352

QY 169 FMIPQYGL-----TLNNGSQAVGRS-----SFVCLF-YFPSQMLRTGNFP 208
Db 353 TKPTQRYLYLSCTRNLPPTTYTGQSQOITDSIQTGLHSDIMFYTIENAVPIHLRTGDEF 412

QY 209 -TPSYTEEVFPFHSSYAHSQSLDR-----LMNPLI--DQVLYLNRTQNSGSAQNK 257
Db 413 STGIYHFDKPL--KLTHSQWTRSLGLPPKLLTEPTTEGDQHPGTLPAASTRKGYHOTI 470

QY 258 DLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIN 317
Db 471 NNSYTEAT-----AIRP-----AQVGYNTPYMNFYSNGGPF-----LT 504

QY 318 P--GTAMASHKDEDEKFFPMSGVMIFGKESAGASNTALDNVMTDEEEIKATNP----- 369
Db 505 PIVPTADTQYNDDE---PNGAIRFTMGVQHGLTTS-----SQELERYTFNPQSKCR 554

QY 370 VATERFGTVA-VNFQSS-----TDPATG--DVHAMGALP-----GMVQWDRDVYLOQP 415
Db 555 APKQFQNAQAPLNLENTNNGTLPSDPIGCKPNMHFMNTLTYGTLTALNNTAPVFNGQ 614

QY 416 IWAKIPIHTD--GHFHPSPLMGGFGLK-NPPPPQILIKNTVPVANPPAEFSA-TKFASFITQ 471
Db 615 IWDKELDTLKPRLH---VTAPFVCKNNPPGQLFVKIAP---NLTDDEFNADSPQQRPIIT 668

QY 472 YSTGQVSVEIEWELQKENSKRNPDEVQYTSNYAKSANVDFVTNNGGLYTEPRPIG 526
Db 669 YSNFWWKGTLTFTAKMRSSNMWNPQQHTT-----TAENIGNYI-PTNIG 712
```

```
RESULT 7
COAT_PAVL3
AC P36310;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Diffoot N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and
RT localization of a unique sequence possibly responsible for its
RT encapsidation pattern.";
RL Virology 192:339-345(1993).
```

```
CC CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81888; -; NOT_ANNOTATED_CDS.
CC PIR; B44276; B44276.
CC HSSP; P07302; 1MVM.
CC InterPro; IPR001403; Parvo_coat.
CC Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 511 511 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 587 AA; 65429 MW; 523E6B9CBF2EBB74 CRC64;

Query Match 7.9%; Score 229; DB 1; Length 587;
Best Local Similarity 22.7%; Pred. No. 4.9e-09;
Matches 133; Conservative 74; Mismatches 228; Indels 150; Gaps 28;

QY 3 SGGGAPMADNNEGADGVGNASGNWHCHDS--TWLGD---RVITSTRTWAL--PTYNN--- 52
Db 30 SGGG-----GSGGGGVGVSTGSYDNQTHYKFLGDGWEITAYSTRMVHLNMPKSENYCR 83

QY 53 ---HLYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSRDMQRLINNNWGRPKR 109
Db 84 VRVHTNTDTGTASHNAMDDAHEQIW-TPWSLVVANAWGVWFQPSDWQYISNNMIHLNHS 142

QY 110 LNFKLFIQVKEVT-TNDGVTTIA---NNLTSTQVFSSEYQLPYVLGSAHQGL--- 161
Db 143 LDQELFNVIKTVTEQNTGAEAIKVYNNDLTAAMVVALDSNNILPYTPAIDNQETLGFYP 202

QY 162 --PPEPA-----DVFMIPOGYLTLNNGSQAVGRSSFCLEYFPS-QMLRTGN 206
Db 203 WKPTIPSPRYFYFSCDRNLSVTYKDEAGTITDTMGLASGLNSOFFTIENTQRINLRTGD 262

QY 207 NF-TFSYTFEEVPHSSYAHQSLSLDRMLNPLDQ-----YLYLNRTQN- 249
Db 263 EYATGTYFDTEPIRLTHWTQNRHLGQPPQITELPSSDTANATLTARGYRSLTQIQGR 322

QY 250 -----QGSQAQNKD-LLFSRGSFAGMSVQPKNWLPG-----PCYRQQR 286
Db 323 NDVTEATRVPRPAQVGFQPHDNFETSIRAGPKFVPVVPADITQGLDHDANGSLRYTDKQH 382

QY 287 VSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMIFGKESA 346
Db 383 GQSWASQNNKDRYTW-DAVNYD-SGR----- 406

QY 347 GASNTALDNVMTDEEEIKATNPVATER---FGTVAVNFQS--SSTDPATGDVHAMGALP 401
Db 407 WTNNCFIQSVPTSEPN---ANQILTNRDNLGAKTDIHFTNAFNSYGPLTAFPH----- 457

QY 402 GMVWQDRDVYLOQPIWAK---IPHTDGHFHPSPLMGGFGLKNPPQILIKNTP--VPANP 456
Db 458 -----PAPIYPOGQIWDKELDLLEHKPLRLHTQAPFV---CKNAPGQLLRLAPNLTDQYD 509

QY 457 PAEFSATKFAFITQYSTGQVSVEIEWELQKENSKRNPDEVQYTS 501
Db 510 PNSSNLSRIVTYGTFFWKGLTLKAK---MRPNA-TWNPVFAQISA 550
```

```
RESULT 8
COAT_PAVHH
ID COAT_PAVHH STANDARD; PRT; 722 AA.
AC P03136;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X01457; CAB57285.1; ALT_SEQ.
DR PIR; A03699; VCPV2.
DR HSSP; P07302; 1MVM.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 722 COAT PROTEIN VP1.
FT CHAIN 131 722 COAT PROTEIN VP2.
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 647 647 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 675 675 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 155 170 GLY-RICH.
SQ SEQUENCE 722 AA; 79737 MW; 6BB678391AA5DC31 CRC64;

Query Match 7.8%; Score 226; DB 1; Length 722;
Best Local Similarity 21.7%; Pred. No. 1.1e-08;
Matches 128; Conservative 83; Mismatches 234; Indels 146; Gaps 26;

QY 2 ASGGAPMADNNEGADGVGNAGSNWHCDSTW--LGDRVITSTRTWAL----- 47
DB 153 ADGGGS-SGGGSGGGIGVSTGTQYDNTQTYKFLGDGWVEITAHASRLHLGMPPSNYC 211
QY 48 --PTYNHLYKQISSASTGASNDNHFGYSTPWGYFDNRFCHFSRDPDQRLINNWGF 105
DB 212 RVTVHNQTTHGTVKGNMAYDTHQIWPWSLVDANAGVWVFPDSDWQFIQNSMESL 270
QY 106 RPKRLNFKLFNIQVKEVTT-----NDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGC 160
DB 271 NLDSLQELFNVVVKTVTEQQAGQDAIKVYVNDLTACMVVALDSNNILPYTPAAQTSET 330
QY 161 L-----PPFPAD---VPMIPQYGLTLNNGSQ-----AVG-----RSSFCYCLE-YFPS 199
DB 331 LGFYPWKPTAPARYYFFMPRQLSVTSSAEGTQITDTIGEPQALNSQFFTIENLTLP 390
QY 200 QMLRTGNFTF-SYTFEEVPFHSSYA-----HSQSLDRLMN-PLIDQYLYLNRTONQSG 252
DB 391 TLLRTGDEFTTGTIYFNTDPLKLTHTWQTNRHACLQGITDLPSTATASLTANGDRFG 450
QY 253 SAQNKOLLF-----SRGSPAGMSVQPKNWLPGPCYRQQRVSK 289
DB 451 STQTQNVYVTEALRTRPAQIGFMQPHDNFEANRGPGFKVPVVP-----LDI 497
QY 290 TKTDNNNSNFTWTGASKYNLNGR--ESIINPGTA-----MASHKDEDEKFFPMSG 337
```

```
DB 498 TAGEDHDAN-----GAIRFNYGKHGEDWAKQGAAPERYTTWDAIDSAAGRDTARCFV---- 549
QY 338 VMIFGESAGASNTALDNVMITDEEEIKA-TNPVATERFGTVAVNFQSSSTDPAIGDVA 396
DB 550 -----QSAPISIPPNQOILQREDAIAGRTNMHYTNVNSYGLSAPFPDP----- 596
QY 397 MGALPGMVQDRDVLQGPWAK---IPHTDGHFHPSPMLMGFGGLKXNPPQILIKNTVPV 453
DB 597 -----IYPNGQIWDKELDEHKPRLHVTAPFV---CKNPPGQLFVHLGP-- 638
QY 454 ANPPAEF--SATKFASFITQYSTGQSVSEIEWELQKENSKRWNPEVQYTSN 502
DB 639 -NLTDQFDPNSTVSRIVT-YSTFYWKGLKFKAKLRPNLTWNVPYQATD 687

RESULT 9
COAT_MUMIV
ID COAT_MUMIV STANDARD; PRT; 716 AA.
AC P03137;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; V01115; CAA24310.1; ALT_SEQ.
DR PIR; A03700; VCPV2M.
DR HSSP; P07302; 1MVM.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 716 COAT PROTEIN VP1.
FT CHAIN 131 716 COAT PROTEIN VP2.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 156 171 GLY-RICH.
SQ SEQUENCE 716 AA; 78707 MW; 6A7229A91161FAC6 CRC64;

Query Match 6.8%; Score 197; DB 1; Length 716;
Best Local Similarity 20.0%; Pred. No. 1.3e-06;
Matches 126; Conservative 88; Mismatches 241; Indels 174; Gaps 28;

QY 3 SGGGAPMADNNEGADGVGNAGSNWHCDSS--TWLGD---RVITTSTRTWALPTYNNHLYKQ 57
DB 160 SGGG-----GSGGGGVGVSTGSDYDQTHYRFLGDGWVEITATRLVHLNMPKSENVC 213
QY 58 I-----SSASTGASNDNHFGYSTPWGYFDNRFCHFSRDPDQRLINNWGFRPKRL 110
DB 214 IRVHTNTDTSVKGNMAKDDAHEQIWTWPSLVDANAWGVWLQPSDWQYICNTMSQLNLVSL 273
QY 111 NFKLFNIQVKEVTTND-----GVTTIANNLSTVQVFSDSYQLPYVLGSAHQGLPPPPA 166
```


Db 274 DQEIFNVVLKTVTEQDLGGQAIKIYNDLTACMVAVDSSNNILFYTPAANSMETLGFYPW 333
Qy 167 DVFMIPQGY-----LTLNNGSQAV-----GRSSPFCLEYFPS-QMLRTGNN 207
Db 334 KPTIASPYRYFCVDRDLSVTYENQEGTVEHNVMTGPKGIPQFFTIENQOITLLRTGDE 393
Qy 208 F-TFSYTFEEVPHSSYAHQSQSLDRLMN--PLIDQYLYLNRTQNSQSGAQNKDLESRG 264
Db 394 FATGTYFDNTSV--KLTHWTQNRQLGQPPLSTF-----PEADTDAGT-----LTAQG 441
Qy 265 SPAGMSVQPKNWL-----PGPCYRQQRVSKTKTDNNSNFT 300
Db 442 SRHGTTQMGVNMVSEAIRTPAQVGFQCPHNDFEASRAGP-FAAPKVPADITQGVDKKAN 500
Qy 301 WTGASKYNLNGRESIINPG-----TAMASHKODEDKPF-----PMSGVMIF 341
Db 501 GSVRYSYGKHQGENWASHGPAPERYTWDETSFGSGRDTKDGFIQSAPLVVPPPLNGI--- 557
Qy 342 GKESAGASNTALDNVMITDBEIKATNPVATERFGTVAVNFQS--SSTDPAATGDVHAMGA 399
Db 558 -----LTNANPIGTKN---DIHFSNVFNSYGPLTAFSH---- 587
Qy 400 LPGMWQDRDVLQGIWAK---IPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTP--VPA 454
Db 588 -----PSPVYPOQIWDKELDLEHKPRLHITAPFV---CKKNAPGQMLVRLGNPLTDQ 637
Qy 455 NPPAEFSATKPFASFITQYSTQGVSVIEIWELOKENSKRWNBEVQYTSNYAKSANVDFTVD 514
Db 638 YDPNGATLSRIVTYGTFFFWKGLTMRACLRA-----NTTNWPNVYQVSAB-----D 682
Qy 515 NNGLY---TEPRPIGT-----RYLTRPL 534
Db 683 NGNSYMSVTKWLPATGNMQSVPLITRPV 711

RESULT 10

COAT_MUMIM STANDARD; PRT; 718 AA.
AC P07302; Q9WMH2; Q9WMH3;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Murine minute virus (strain MVM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 132-718.
RA Llamas-Saiz A.L., Agbandje-Mckenna M., Wikoff W.R., Bratton J.,
RA Tattersall P., Rossmann M.G.;
RT "Structure determination of Minute Virus of mice.";
RL Acta Crystallogr. D 53:93-100(1997).
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; CAB46507.1; -.
DR EMBL; X02481; CAB46508.1; -.
DR EMBL; M12032; AAA69569.1; ALT_INIT.
DR PIR; B23008; VCPVIM.
DR PDB; 1MVM; 25-FEB-98.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein; 3D-structure.
FT CHAIN 1 718 COAT PROTEIN VP1.
FT CHAIN 132 718 COAT PROTEIN VP2.
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DOMAIN 157 172 GLY-RICH.
FT CONFLICT 144 144 A -> G (IN REF. 2).
FT STRAND 182 186
FT STRAND 191 204
FT STRAND 213 217
FT TURN 220 221
FT TURN 228 231
FT STRAND 235 245
FT TURN 251 253
FT HELIX 256 265
FT STRAND 266 266
FT STRAND 268 269
FT STRAND 274 289
FT STRAND 296 301
FT TURN 303 304
FT STRAND 307 307
FT STRAND 309 312
FT HELIX 322 325
FT TURN 326 326
FT TURN 333 334
FT STRAND 337 338
FT STRAND 340 345
FT STRAND 351 351
FT TURN 357 358
FT STRAND 367 367
FT TURN 373 375
FT HELIX 381 384
FT STRAND 390 390
FT STRAND 396 397
FT STRAND 408 409
FT STRAND 428 428
FT STRAND 438 438
FT HELIX 442 444
FT TURN 458 460
FT STRAND 466 467
FT TURN 473 474
FT STRAND 477 477
FT STRAND 484 484
FT TURN 494 496
FT TURN 499 503
FT STRAND 509 509
FT TURN 512 513
FT STRAND 525 525
FT HELIX 534 536
FT STRAND 546 546
FT TURN 557 558
FT STRAND 559 559
FT TURN 575 577
FT TURN 583 584
FT STRAND 585 586
FT STRAND 600 600
FT STRAND 618 618
FT STRAND 627 630


```
Db 661 TYSDFWVKGLVFKAKLRASHWTNPIQQMSIN-----VDNQFNYL-PNNIG 705

RESULT 12
COAT_FPV
ID -COAT_FPV STANDARD; PRT; 727 AA.
AC P04864; Q65112;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP MEDLINE=85265017; PubMed=2991581;
RA Carlson J., Rushlow K., Maxwell I., Maxwell F., Winston S., Hahn W.;
RT "Cloning and sequence of DNA encoding structural proteins of the
RT autonomous parvovirus feline panleukopenia virus.";
RL J. Virol. 55:574-587(1985).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10824; AAA47161.1; -
DR EMBL; M10824; AAA47162.1; -
DR FIP; A03701; VCPV1F.
DR HSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
FT CHAIN 1 727 COAT PROTEIN VP1.
FT CHAIN 144 727 COAT PROTEIN VP2.
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 660 660 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 165 180 GLY-RICH.
SQ SEQUENCE 727 AA; 80344 MW; 77E0F6FE554C0C6E CRC64;

Query Match 6.7%; Score 194; DB 1; Length 727;
Best Local Similarity 20.7%; Pred. No. 2.2e-06;
Matches 136; Conservative 73; Mismatches 215; Indels 232; Gaps 32;

QY 5 GGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITST-----RTWALPTYNN- 52
Db 153 GGQP-AVRNERATGSGNGSGGGGGGS--GGVGISTGTFFNNQTEKFLENGWVEITANSS 209
QY 53 ---HL-----YKQI-----SSASTGASNDNHFGYSTFWGVDFENRHFCHFSRDN 95
Db 210 RLVLNMPSENYKRVVNVNMDKTAVKGNMADDDTHVQIVTFWSLVDANAWGVWFNPGDW 269
QY 96 QRLINNNGFRPKRLNFKLNIQVKEVT---TNDGVTTIANLSTVQVFSDEYQLPYV 152
Db 270 QLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT 329
QY 153 LGSAHQGL-----PPFPA-----DVFMIPQYGLTLNNGSQAVGRS----- 189
Db 330 PAAMRSETLGFYPWKPTIPTPRYYFQWDRTLIPSH-----TGTSPTNIYHGTDPDD 383
QY 190 -SFYCLE-YFPSQMLRTGNF---TF-----SYTFEE-----VPEHSSYAHQSLS 229
Db 384 VQFYTIENSVPVHLRLTGDDEFATGTFEEDCKPCRLTHTWTQNRALGLPFFFLNSLPQSEGA 443
```

```
QY 230 DR-----LMNPLIDQYL--YYLNRTQNQS----- 251
Db 444 TNFGDIGVQDDKRRGVTQMGNTDYITEATIMRPAEVGYSAPYTSFEASTQGPFKIPAAAG 503
QY 252 -GSAQNKDLLFSRGSPPAGMSVQPKNWLPGPCY---RQQRVSKTKTDNNNSNFTWTGASKY 307
Db 504 RGAQTDENQAADGDP-----YAFGRHQHGQKTTTGETPERFTY----- 543
QY 308 NLNGRESIINPGTAMASHKDDKFFPMGSMVIFGKESAGASNTALD-NVMITDEEEIKA 366
Db 544 -----IAHQDT-----GRYPAGDWIQNINFNLPVTNDNVLLP 575
QY 367 TNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPVWAKIPHTDGH 426
Db 576 TDPIG---GKTGINY--TNIFNYGPLTALNNVP-----PVYPNGQIWDKEFTD-- 620
QY 427 FHPSPLMGGEGLK-----NPPQILIKNTPVPAN---PPAEFSATKFAFIT 470
Db 621 -----LKPRHLHVNAFVQCNNCPGQLFVKVAPNLNEYDPDASANMSR-----IV 665
QY 471 QYSTQVSVEIEWELQKENSXRNPEVOYTSNYAKSANVDFTVDNNGLYTEPRPIG 526
Db 666 TYSDFWVKGLVFKAKLRASHWTNPIQQMSIN-----VDNQFNYL-PNNIG 710

RESULT 13
COAT_PAVC2
ID -COAT_PAVC2 STANDARD; PRT; 584 AA.
AC P30129;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein VP2.
OS Canine parvovirus (type 2 / strain A72) (CPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=31597;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (3.25 ANGSTROMS).
RX MEDLINE=91173306; PubMed=2006420;
RA Tsao J., Chapman M.S., Agbandje M., Keller W., Smith K., Wu H.,
RA Luo M., Smith T.J., Rossmann M.G., Compans R.W., Parish C.R.;
RT "The three-dimensional structure of canine parvovirus and its
RT functional implications."
RL Science 251:1456-1464(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC OF A COMBINATION OF VP2, VP3, AND SOME VP1.
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC PDB; 4DPV; 01-APR-97.
DR PDB; 1IJS; 23-DEC-96.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein; 3D-structure.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TURN 38 39
FT STRAND 49 53
FT STRAND 58 64
FT STRAND 67 71
FT STRAND 80 84
FT STRAND 87 89
FT STRAND 102 112
FT TURN 118 120
FT HELIX 123 132
FT STRAND 133 156
FT STRAND 162 167
FT TURN 169 170
FT STRAND 173 177
```



```
FT TURN 179 180
FT TURN 188 191
FT TURN 199 200
FT STRAND 203 203
FT STRAND 206 211
FT STRAND 215 218
FT STRAND 231 235
FT TURN 238 239
FT HELIX 246 249
FT STRAND 252 255
FT STRAND 261 262
FT STRAND 266 267
FT STRAND 270 270
FT STRAND 273 274
FT TURN 300 301
FT HELIX 310 312
FT STRAND 317 317
FT TURN 327 329
FT STRAND 330 330
FT STRAND 335 336
FT STRAND 339 339
FT STRAND 346 347
FT STRAND 352 353
FT TURN 365 366
FT TURN 370 374
FT STRAND 378 380
FT TURN 388 389
FT STRAND 396 398
FT TURN 410 411
FT STRAND 415 416
FT TURN 418 419
FT HELIX 426 428
FT STRAND 429 430
FT TURN 445 447
FT STRAND 450 450
FT TURN 453 454
FT STRAND 455 456
FT STRAND 464 464
FT TURN 465 466
FT STRAND 470 470
FT STRAND 488 488
FT STRAND 490 490
FT STRAND 498 502
FT STRAND 506 506
FT TURN 512 513
FT STRAND 520 520
FT STRAND 523 539
FT TURN 555 557
FT HELIX 558 560
FT TURN 565 566
FT STRAND 574 574
SQ SEQUENCE 584 AA; 64686 MW; 6E4DADA5AEBF9D8C CRC64;

Query Match 6.5%; Score 190; DB 1; Length 584;
Best Local Similarity 20.3%; Pred. No. 3.1e-06;
Matches 130; Conservative 82; Mismatches 229; Indels 198; Gaps 30;

Qy 5 GGAPNADNNEGADGVGNASGNMHCDSTWLGDRVITTST-----RTWALPTVNN- 52
Db 10 GGQP-AVRNERATSGNGSGGGGGGS--GGVGISTGTENNQTEFKFLENGWVYITANSS 66
Qy 53 ---HL-----YKQI-----SSASTGASNDNHFGYSTPWGYFDFNRFHCHFSRDLW 95
Db 67 RLVELNMPESENYRRVVVNNMDKTAVNGNMALDDIHAEIVTFWLSLVDANAGWVWFNPGDW 126
Qy 96 QRLINNNGWFRPKRLNFKLEFNIQKVT--TNDGVTTIANLNTSTVQVSDSEYQLPYV 152
Db 127 QLIVNTMSLHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT 186
Qy 153 LGSAHQGL-----PPFPA-----DVFMIPQGYLTLNNGSQAVGRS----- 189
Db 187 PAAMRSETLGFYPWKPTIPTPRYFQWDRTLIPSH-----TGTSGTPTNIYHGTDPPD 240
```

```
Qy 190 -SFYCLE-YFPSQMLRTGNF-TFSYTFEEVPHSSYAHQSOLDRLMN--PLIDQYLYL 244
Db 241 VQYTIENSVPVHLLRTGDEFATGTFDFCKP--CRLTHWTQTRALGLPP-----FL 291
Qy 245 NRTQNSGSAQNKDLLFSRSPAGMSVQPKWLPGPCYRQRVSKTKTDNNNSNF----- 299
Db 292 NSLPOSEGATNFGDI-----GV-----QDKRGRVQTOMGNTNYITEAT 329
Qy 300 ----TWTGASKYNLNGRESIINP-----GTAMASHKDDKFPFMSGVMIFGESAGASN 350
Db 330 IMRPAEVGYSAPYYSFEASTQGPKTPIAAGRGAQTDEQAADGNPRYAFGRHGKKT 389
Qy 351 TALD-----NMVITDEEIKATNPVATERFGTVAVNFQ 383
Db 390 TTGETPERFTYIAHQDTGRYPEGDWIQNINFLPVTNDNVLLPTDPIG-----GKTGINY- 444
Qy 384 SSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLK---- 439
Db 445 -TNIFNTYGPLTALNNVP-----PVYPNGQIWDKEFTD-----LKPRLH 483
Qy 440 -----NPPPQILIKNTPVPAN---PPAEFSATKPFASFITQYSTQVSVIEWELQK 487
Db 484 VNAPFVCQNNCPGQLFVKVAPNLTNQYDPPDASANMSR-----IVTYSDFWVKGLVFKAKL 539
Qy 488 ENSKRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 526
Db 540 RASHTWNPIQMSIN-----VDNQFNIV-PSNIG 567

RESULT 14
COAT_FPV19
ID_COAT_FPV19 STANDARD; PRT; 727 AA.
AC P24840;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein vp1 [Contains: Coat protein VP2].
OS Feline panleukopenia virus (strain 193) (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=193/70;
RX MEDLINE=91073139; PubMed=2174965;
RA Martyn J.C., Davidson B.E., Studdert M.J.;
RT "Nucleotide sequence of feline panleukopenia virus: comparison with
canine parvovirus identifies host-specific differences.";
RL J. Gen. Virol. 71:2747-2753(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CU-4;
RX MEDLINE=91272479; PubMed=1647068;
RA Parrish C.R.;
RT "Mapping specific functions in the capsid structure of canine
parvovirus and feline panleukopenia virus using infectious plasmid
clones.";
RL Virology 183:195-205(1991).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X55115; CAA38911.1; --
DR EMBL; M38246; AAC37928.1; --
DR EMBL; M38246; AAC37929.1; --
DR PIR; B36608; VCPVFP.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
```

DR	Pfam; PF00740; Parvo_coat; 1.	OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
KW	Coat protein; Glycoprotein.	OX	NCBI_TaxID=59284;
FT	CHAIN 1 727	RN	[1]
FT	CHAIN 144 727	RX	SEQUENCE FROM N.A.
FT	DOMAIN 165 190	RA	MEDLINE=91272479; PubMed=1647068;
FT	CARBOHYD 168 168	RT	Parrish C.R.;
FT	CARBOHYD 190 190	RT	"Mapping specific functions in the capsid structure of canine
FT	CARBOHYD 207 207	RT	parvovirus and feline panleukopenia virus using infectious plasmid
FT	CARBOHYD 323 323	RT	clones.";
FT	CARBOHYD 586 586	RL	Virology 183:195-205(1991).
FT	CARBOHYD 648 648	CC	-!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
FT	CARBOHYD 660 660	CC	-!- OF A COMBINATION OF VP2, VP3, AND SOME VP1.
FT	DOMAIN 165 180	CC	-!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
SQ	SEQUENCE 727 AA; 80386 MW; 648596C09B621FF5 CRC64;	CC	-----
Query Match 6.5%; Score 189; DB 1; Length 727;		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
Best Local Similarity 20.6%; Pred. No. 5e-06;		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
Matches 135; Conservative 73; Mismatches 216; Indels 232; Gaps 32;		CC	the European Bioinformatics Institute. There are no restrictions on its
		CC	use by non-profit institutions as long as its content is in no way
		CC	modified and this statement is not removed. Usage by and for commercial
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
		CC	or send an email to license@isb-sib.ch).
		CC	-----
QY	5 GGAPMADNNEGADVGNASGNWHCDSTWLGDRVTTST-----RTWALPTYNN- 52	DR	EMBL; M38245; AAB02799.1; -.
Db	153 GGQP-AVRNERATSGNGSGGGGGGS--GGVGISTGTENNQTBFKFLENGWVEITANSS 209	DR	EMBL; M38245; AAB02800.1; -.
QY	53 ---HL-----YKQI-----SSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDW 95	DR	HSSP; P30129; 4DPV.
Db	210 RLVLNMPESSENYKRVVNNMMDKTAVKGNMALLDHIHQIVTPWSLVDANAWGVFNPGDW 269	DR	InterPro; IPR001403; Parvo_coat.
QY	96 QRLINNNGFRPKRLNFKLFNIQVKEVT---TNDGVTTIANNLSTVQVFSSEYQLPYV 152	KW	Coat protein; Glycoprotein.
Db	270 QLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT 329	FT	CHAIN 1 727
QY	153 LGSAHQGL-----PPFPA-----DVFMIPQGYGLTLNNGSQAVGRS----- 189	FT	CHAIN 144 727
Db	330 PAAMRSETLGFYPWKPTIPTPWRYFQWDRTLIPSH-----TGTSGTPTNVYHGTDPDD 383	FT	CARBOHYD 168 168
QY	190 -SFYCLE-YFPSQMLRTGNF---TF-----SYTFEE-----VPFHSSYAHQSQSL 229	FT	CARBOHYD 190 190
Db	384 VQFYTIENSVPVHLRLRTGDEFATGTFPFDCKPCRLTHTWQTNRALGLPPFLNSLPQSEGA 443	FT	CARBOHYD 207 207
QY	230 DR-----LMNPLIDQYL--YYLNRTQNS----- 251	FT	CARBOHYD 323 323
Db	444 TNFGDIGVQDQKRRGVTQMGNTDYTEATIMRPAEVGYSAPYSEASTQGPFKTPIAAG 503	FT	CARBOHYD 586 586
QY	252 -GSAQNKDLLFSRGSPAGMSVQPKNWLPGPCY---RQQRVSKTKTDNNNSNFTWTGASKY 307	FT	CARBOHYD 648 648
Db	504 RGAQTDENQAADGDP-----YAFGRHQGQKTTTGETPERFTY----- 543	FT	CARBOHYD 660 660
QY	308 NLNGRESINPGTAMASHKDDKEDKFFPMGSMVIFGKESAGASNTALD-NVMITDEEEIKA 366	SQ	SEQUENCE 727 AA; 80342 MW; 115F3E1A79098EBE CRC64;
Db	544 -----IAHQDT-----GRYPEGDIQINFNFLPVTNDNVLLP 575	Query Match 6.5%; Score 188; DB 1; Length 727;	
QY	367 TNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKIPHDTGH 426	Best Local Similarity 20.3%; Pred. No. 5.9e-06;	
Db	576 TDPIG-----GKTGINY--TNIFNTYGLTALNNVP-----PVYPNGQIWDKEFDTD-- 620	Matches 130; Conservative 82; Mismatches 229; Indels 198; Gaps 30;	
QY	427 FHPSPLMGGFGLK-----NPPQILIKNTVPVPA-----PPAEFSATKFASFIT 470	QY	5 GGAPMADNNEGADVGNASGNWHCDSTWLGDRVTTST-----RTWALPTYNN- 52
Db	621 -----LKPRLHVNAFPVQNNCPGQLFVKVAPNLITNEYDPPDASANMSR-----IV 665	Db	153 GGQP-AVRNERATSGNGSGGGGGGS--GGVGISTGTENNQTBFKFLENGWVEITANSS 209
QY	471 QYSTGVSVIEIWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 526	QY	53 ---HL-----YKQI-----SSASTGASNDNHFGYSTPWGYFDNRFCHFSPRDW 95
Db	666 TYSDFWKGKLVFKAKLRASHTWNPQQMSIN-----VDNQFNIV-PNNIG 710	Db	210 RLVLNMPESSENYKRVVNNMMDKTAVKGNMALLDHIHQIVTPWSLVDANAWGVFNPGDW 269
RESULT 15		QY	96 QRLINNNGFRPKRLNFKLFNIQVKEVT---TNDGVTTIANNLSTVQVFSSEYQLPYV 152
COAT_PAVCB		Db	270 QLIVNTMSELHLVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNTMPFT 329
ID	STANDARD; PRT; 727 AA.	QY	153 LGSAHQGL-----PPFPA-----DVFMIPQGYGLTLNNGSQAVGRS----- 189
AC	Q11213;	Db	330 PAAMRSETLGFYPWKPTIPTPWRYFQWDRTLIPSH-----TGTSGTPTNVYHGTDPDD 383
DT	01-OCT-1996 (Rel. 34, Created)	QY	190 -SFYCLE-YFPSQMLRTGNF---TF-----SYTFEE-----VPFHSSYAHQSQSL 229
DT	01-OCT-1996 (Rel. 34, Last sequence update)	Db	384 VQFYTIENSVPVHLRLRTGDEFATGTFPFDCKPCRLTHTWQTNRALGLPPFLNSLPQSEGA 443
DT	16-OCT-2001 (Rel. 40, Last annotation update)	QY	230 DR-----LMNPLIDQYL--YYLNRTQNS----- 251
DE	Coat protein VP1 [Contains: Coat protein VP2].	Db	444 TNFGDIGVQDQKRRGVTQMGNTDYTEATIMRPAEVGYSAPYSEASTQGPFKTPIAAG 503
OS	Canine parvovirus (strain B) (CPV).	QY	252 -GSAQNKDLLFSRGSPAGMSVQPKNWLPGPCY---RQQRVSKTKTDNNNSNFTWTGASKY 307

Db 533 TTGETPERFTYIAHQDTGRYPEGDWIQINFNLPVTNDNVLLPTDPIG-----GKTGINY- 587

Qy 384 SSSTDPAATGDVHAMGALPGMWQDRDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLK----- 439

Db 588 -TNIFNTYGPLTALNNVP-----PVYPNGQIWDKEFDTD-----LKPRLH 626

Qy 440 -----NPPPPQILIKNTPVPAN---PPAEFSATKPFASFITQYSTGQVSVIEWELQK 487

Db 627 VNAPFVCQNNCPGQLFVKVAPNLTNEYDPPDASANMSR-----IVTYSDFWWKGLVFKAKL 682

Qy 488 ENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIG 526

Db 683 RASHTWNPQQMSIN-----VDNQFNYYV-PSNIG 710

Search completed: January 21, 2004, 16:02:14
Job time : 11.1429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:00:06 ; Search time 27.7143 Seconds
(without alignments)
4972.168 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				SPTREMBL 23:*			
				1:	sp_archaea:*		
				2:	sp_bacteria:*		
				3:	sp_fungi:*		
				4:	sp_human:*		
				5:	sp_invertebrate:*		
				6:	sp_mammal:*		
				7:	sp_mhc:*		
				8:	sp_organelle:*		
				9:	sp_phage:*		
				10:	sp_plant:*		
				11:	sp_rodent:*		
				12:	sp_virus:*		
				13:	sp_vertebrate:*		
				14:	sp_unclassified:*		
				15:	sp_rvirus:*		
				16:	sp_bacteriap:*		
				17:	sp_archaeap:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.		Score	Query Match	Length	DB ID	Description
SUMMARIES						
1	2906	100.0	736	12	Q9WBP8	Q9wbp8 adeno-assoc
2	2884	99.2	736	12	O56137	O56137 adeno-assoc
3	2530	87.1	736	12	O56139	O56139 adeno-assoc
4	2528	87.0	736	12	Q65311	Q65311 adeno-assoc
5	2481.5	85.4	533	12	O92917	O92917 adeno-assoc
6	2481.5	85.4	598	12	O56653	O56653 adeno-assoc
7	2481.5	85.4	735	12	O56652	O56652 adeno-assoc
8	2430	83.6	737	12	Q8JQG0	Q8jqg0 adeno-assoc
9	2384.5	82.1	738	12	Q8JQF8	Q8jqf8 adeno-assoc
10	1698	58.4	534	12	Q67668	Q67668 goose parvo
11	1698	58.4	587	12	Q67667	Q67667 goose parvo
12	1698	58.4	732	12	Q67666	Q67666 goose parvo
13	1697	58.4	732	12	Q8V395	Q8v395 goose parvo
14	1690.5	58.2	734	12	O41855	O41855 adeno-assoc
15	1690	58.2	732	12	Q83290	Q83290 muscovy duc
16	1676	57.7	534	12	Q65446	Q65446 barbarie du

17	1676	57.7	587	12	Q65445	Q65445 barbarie du
18	1676	57.7	732	12	Q65444	Q65444 barbarie du
19	1676	57.7	732	12	Q83289	Q83289 muscovy duc
20	1665	57.3	724	12	Q9YIJ1	Q9yij1 adeno-assoc
21	1521	52.3	676	12	Q67672	Q67672 goose parvo
22	715	24.6	179	12	Q9W8U7	Q9w8u7 duck parvov
23	715	24.6	179	12	Q9WA24	Q9wa24 goose parvo
24	711	24.5	179	12	Q9WN18	Q9wn18 duck parvov
25	709	24.4	179	12	Q9WN19	Q9wn19 duck parvov
26	707	24.3	179	12	Q9WN20	Q9wn20 goose parvo
27	614	21.1	571	12	Q8QQV4	Q8qqv4 minute viru
28	614	21.1	703	12	Q8QQV5	Q8qqv5 minute viru
29	536.5	18.5	947	12	Q918U9	Q918u9 bovine parv
30	501	17.2	554	12	Q9PZS9	Q9pzs9 human parvo
31	501	17.2	781	12	Q9PZT0	Q9pzt0 human parvo
32	496	17.1	554	12	Q8JN54	Q8jns4 human parvo
33	496	17.1	781	12	Q8JN56	Q8jns6 human parvo
34	495	17.0	785	12	Q9J0X4	Q9j0x4 pig-tailed
35	494.5	17.0	554	12	Q9JGP7	Q9jgp7 human parvo
36	494.5	17.0	781	12	Q9JGP8	Q9jgp8 human parvo
37	493	17.0	546	12	Q913X0	Q913x0 human parvo
38	493	17.0	773	12	Q913X1	Q913x1 human parvo
39	493	17.0	781	12	Q8JYE3	Q8jye3 erythroviru
40	492.5	16.9	781	12	P89318	P89318 human parvo
41	492.5	16.9	781	12	P89317	P89317 human parvo
42	492	16.9	781	12	P89319	P89319 human parvo
43	492	16.9	781	12	Q8JYD9	Q8jyd9 erythroviru
44	490.5	16.9	781	12	P89316	P89316 human parvo
45	490	16.9	554	12	Q90201	Q90201 human parvo

ALIGNMENTS

RESULT 1

Q9WBP8 ID Q9WBP8 PRELIMINARY; PRT; 736 AA.

AC Q9WBP8;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE Capsid protein.

OS Adeno-associated virus 1.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

OX NCBI_TaxID=85106;

RN [1]_TaxID=85106;

RP SEQUENCE FROM N.A.

RX MEDLINE=99214338; PubMed=10196295;

RA Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;

RT "Gene therapy vectors based on adeno-associated virus type 1.";

RL J. Virol. 73:3994-4003(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Xiao W., Wilson J.M.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF063497; AAD27757.1; -.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

SQ SEQUENCE 736 AA; 81375 MW; CFABFB9BD5CD0595 CRC64;

Query Match 100.0%; Score 2906; DB 12; Length 736;

Best Local Similarity 100.0%; Pred. No. 8.3e-208;

Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 60
|||||

Db 203 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISS 262
|||||

QY 61 ASTGASNDNHFGYSTPWGYDFDNRFCHFSPRDQRLNNNWGFRPKRLNFKLFNIQVK 120
|||||

Db 263 ASTGASNDNHFGYSTPWGYDFDNRFCHFSPRDQRLNNNWGFRPKRLNFKLFNIQVK 322
|||||

QY 121 EVTTNDGVTTIANNLTSTVQFSDSEYQLPYVLGSAHQCLPPFPADVFEMIPQYGYLTN 180
|||||

Db 323 EVTTNDGVTTIANNLTSITVQFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 382
Qy 181 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMPLIDQY 240
Db 383 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMPLIDQY 442
Qy 241 LYTLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 300
Db 443 LYTLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGVMI FGKESAGASNTALDNVMITD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGVMI FGKESAGASNTALDNVMITD 562
Qy 361 EEEIKATNPVATERFGTVAVNPFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKI 420
Db 563 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGVMI FGKESAGASNTALDNVMITD 562
Qy 421 EEEIKATNPVATERFGTVAVNPFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKI 420
Db 563 EEEIKATNPVATERFGTVAVNPFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKI 622
Qy 421 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 623 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 682
Qy 481 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
Db 683 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 736

RESULT 2
O56137
ID O56137 PRELIMINARY; PRT; 736 AA.
AC O56137;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028704; AAB95450.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565P5 CRC64;

Query Match 99.2%; Score 2884; DB 12; Length 736;
Best Local Similarity 99.1%; Pred. No. 3.6e-206;
Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 262
Qy 61 ASTGASNDNHYFGYSTPWGYFDNRFCHFS PRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 263 ASTGASNDNHYFGYSTPWGYFDNRFCHFS PRDWQRLINNNWGFPRKRLNFKLFNIQVK 322
Qy 121 EVTTNDGVTTIANNLTSITVQFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 180
Db 323 EVTTNDGVTTIANNLTSITVQFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 382
Qy 181 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMPLIDQY 240

Db 383 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSQSLDRLMPLIDQY 442
Qy 241 LYTLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 300
Db 443 LYTLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGVMI FGKESAGASNTALDNVMITD 360
Db 503 WTGASKYNLNGRESIINPGTAMASHKDDKDEKFFPMSGVMI FGKESAGASNTALDNVMITD 562
Qy 361 EEEIKATNPVATERFGTVAVNPFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKI 420
Db 563 EEEIKATNPVATERFGTVAVNLFQSSSTDPATGDVHMVGMALPGMVWQDRDVYLQGPWAKI 622
Qy 421 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 623 PHTDGHFHPSPMLGGFGLKHPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 682
Qy 481 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
Db 683 IEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 736

RESULT 3
O56139
ID O56139 PRELIMINARY; PRT; 736 AA.
AC O56139;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Capsid protein VP1.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028705; AAB95452.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81906 MW; DD52331AD5F0D70F CRC64;

Query Match 87.1%; Score 2530; DB 12; Length 736;
Best Local Similarity 86.0%; Pred. No. 8.2e-180;
Matches 460; Conservative 27; Mismatches 46; Indels 2; Gaps 2;
Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVTTSTRTWALPTYNNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISS 262
Qy 61 ASTGASNDNHYFGYSTPWGYFDNRFCHFS PRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHYFGYSTPWGYFDNRFCHFS PRDWQRLINNNWGFPRKLSFKLFNIQVK 321
Qy 121 EVTTNDGVTTIANNLTSITVQFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTIN 180
Db 322 EVTQNDGTTTIANNLTSITVQFTDSEYQLPYVLGSAHQGCLPPFPADVFMVFPQYGYLTIN 381
Qy 181 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMPLIDQY 240
Db 382 NGSQAVGRSSFCYCLEYFPPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMPLIDQY 441
Qy 241 LYTLNRTQ-NSGSAQNKDLLFSRGS PAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNF 299
Db 442 LYTLNRTQGTSGTTNQSRLLFSQAGPQSMQSLQARNWLPGPCYRQORLSKTANDNNNSNF 501


```
QY 301 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNVMITD 360
Db 300 WTGATKYHLNGRDSLVPNPGPAMASHKODEEKFPPQSGVLIFGKQSEKTNVDIEKVMITD 359
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 360 EEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVQDRDVLQGPWAKI 419
QY 421 PHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 420 PHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPSTTFSAKFASFITQYSTGQVSVE 479
QY 481 IEWELQKENSKRWNPEVOYQTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 480 IEWELQKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 533

RESULT 6
O56653 PRELIMINARY; PRT; 598 AA.
AC O56653;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Major coat protein VP2.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
RT proteins affect viral infectivity: lack of an RGD integrin-binding
RT motif.";
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03778.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
FT VARIANT 416 416 D -> N.
FT VARIANT 430 430 T -> N.
FT VARIANT 540 541 QV -> HV.
FT VARIANT 573 573 V -> R.
SQ SEQUENCE 598 AA; 66619 MW; 070811ED9368E934 CRC64;

Query Match 85.4%; Score 2481.5; DB 12; Length 598;
Best Local Similarity 83.3%; Pred. No. 2.5e-176;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 66 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLYKQISS 125
QY 61 ASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 126 QS-GASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 184
QY 121 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTIN 180
Db 185 EVTQNDGVTTIANNLTSTVQVFTDSYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTIN 244
QY 181 NGSQAVGRSSFCLEYFPQSOMLRTGNFTFSYTFEEVPPHSSVAHSQSGLDRMLNPLIDQY 240
Db 245 NGSQAVGRSSFCLEYFPQSOMLRTGNFTFSYTFEDVPPHSSVAHSQSGLDRMLNPLIDQY 304
```

```
QY 241 LYLLNRTQNSGAQNKKOLLFSRSPAGMSVQPKNWLPGCYRQORVSKTKTDNNNSFT 300
Db 305 LYLLSRITNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGCYRQORVSKTSADNNNSEYS 364
QY 301 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMIFGKESAGASNTALDNVMITD 360
Db 365 WTGATKYHLNGRDSLVPNPGPAMASHKODEEKFPPQSGVLIFGKQSEKTNVDIEKVMITD 424
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 425 EEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVQDRDVLQGPWAKI 484
QY 421 PHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 485 PHTDGHFHPSPMLMGFGFLKNPPPPQILIKNTVPANPPSTTFSAKFASFITQYSTGQVSVE 544
QY 481 IEWELQKENSKRWNPEVOYQTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 545 IEWELQKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 7
O56652 PRELIMINARY; PRT; 735 AA.
AC O56652;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Major coat protein VP1.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
RT proteins affect viral infectivity: lack of an RGD integrin-binding
RT motif.";
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03780.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
FT VARIANT 76 76 D -> V.
FT VARIANT 553 553 D -> N.
FT VARIANT 567 567 T -> N.
FT VARIANT 677 678 QV -> HV.
FT VARIANT 710 710 V -> R.
SQ SEQUENCE 735 AA; 81944 MW; 980BEEF469083908 CRC64;

Query Match 85.4%; Score 2481.5; DB 12; Length 735;
Best Local Similarity 83.3%; Pred. No. 3.3e-176;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNNHLYKQISS 262
QY 61 ASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHYFGYSTPWGYFDENRFCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 321
QY 121 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTIN 180
Db 322 EVTQNDGVTTIANNLTSTVQVFTDSYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTIN 381
```

QY	181	NGSQAVGRSSFYCLEYFPFSQMLRTGNNFTFSYTTFEEVPHSSHSAHSQSGLDRLMNPLIDQY	240
Dd	382	NGSQAVGRSSFYCLEYFPFSQMLRTGNNFTFSYTTFEDVPFHSSHSAHSQSGLDRLMNPLIDQY	441
QY	241	LYYLNRNTQNQSGSAQNKDLLFSGSPAGMSVPQKNWLPGPCYRQQRVSKTCTDNNNSNFT	300
Dd	442	LYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGPCYRQQRVSKTADNNNSEYS	501
QY	301	WTGASKYNLNGRESIIINPGTAMASHKDDEDKFFPMSGVMIFGKESAGASNATALDNVMITD	360
Dd	502	WTGATKYHLNGRDSLVPNPGPAMASHKDDEKEFFPQSGVLIFGKGQSEKTVNDIEKVMITD	561
QY	361	EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVYLOGPIWAKI	420
Dd	562	EEEIRTTNPVATEQYGSVSTNLQRGNRQAATAADVNTQGVLPGMWQDRDVYLOGPIWAKI	621
QY	421	PHTDGHFHPSPLMGGFGLKNPPQILIKNTVPANPPAEFSATKKFASFITYOSTGOVSVE	480
Dd	622	PHTDGHFHPSPLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITYOSTGOVSVE	681
QY	481	IIEWELOKENSKRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTLYLTRPL	534
Dd	682	IIEWELOKENSKRWNPETQYTSNYKNSVNVDFTVDNTGVSEPRPIGTLYLTRNL	735

RESULT 8

Q8JQG0 PRELIMINARY; PRT; 737 AA.

Query Match	83.6%	Score 2430;	DB 12;	Length 737;
Best Local Similarity	82.1%;	Pred. No. 2.3e-172;		
Matches 440;	Conservative 45;	Mismatches 47;	Indels 4;	Gaps 3;

Qy	1	MASGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS	60
Db	204	VAAGGGAPMADNNEGADGVGNASGNWHCHDSTWLGDRVITTTSTRTWALPTYNHLYKQISS	263
Qy	61	ASTGASNDNHFGYSTPWGYFDENRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVK	120
Db	264	ETAGSTNDNTYFGYSTPWGYFDENRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVK	323
Qy	121	EVTTNDGVTTIANNLSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTNL	180
Db	324	EVTTNDGVTTIANNLSTIQVFSDSYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTNL	383
Qy	181	NGSQAVGRSSFYCLEYFPPSOMLRTGNNTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQY	240
Db	384	NGSQSVGRSSFYCLEYFPPSQMLRTGNNFESYSFEDVPFHSSYAHSQSLDRLMNPLIDQY	443
Qy	241	LYYLNRTQ-NQSGSAQNKDLLFSPRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNF	299
Db	444	LYYLARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNF	503
Qy	300	TWTGASKYNLNGRESIINPGTAMASHKDDDEDKFFPMMSGVMI FGKESAGASN - TALDNTVMI	358

Db	504	AWTGATKYHLGNRSLVNFQVAMATHKDDERFPSSGVLIIFGK--TGATNKTTLNVLM	561
QY	359	TDEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDQVYLGQPIWA	418
Db	562	TNEEEIRPTNPVATEEYGIYSSNLQAANTAQTQVVNNQGALPGMVWQNRDQVYLGQPIWA	621
QY	419	KIPHTDGHFHPSPLMGGFGLKNPPPPQILIKNTVPVANPAAEFSA TKFASFITQYSTGQVS	478
Db	622	KIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPVANPPEVFTPAKFA SFITQYSTGQVS	681
QY	479	VEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL	534
Db	682	VEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL	737

RESULT 9

Q8JQF8
ID Q8JQF8 PRELIMINARY; PRT; 738 AA.

Query Match 82.1%; Score 2384.5; DB 12; Length 738;
Best Local Similarity 80.2%; Pred. No. 5.6e-169;
Matches 429; Conservative 44; Mismatches 61; Indels 1; Gaps 1

QY	1	MASGGGAPMADNNEGADVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS	60
Db	204	MAAGGGAPMADNNEGADVGVSSSGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISN	263
QY	61	-ASTGASNDNHFGYSTPWGYFDENRFHCHFSPRDWORLINNNWGFPRKRLNFKLENIQV	119
Db	264	GTSGGATNDNTYFGYSTPWGYFDENRFHCHFSPRDWORLINNNWGFPRKRLSPKLFNIQV	323
QY	120	KEVTTNDGVTTIANNLTTSTVQVPSDSEYQLPYVLSAHQGLPPFPADVFMIPQYGYLTL	179
Db	324	KEVTQNEGKTIANNLTSTIOVFTDSEYQLPYVLSAHQGLPPFPADVFMIPQYGYLTL	383
QY	180	NNGSOAVGRSSPYCLEYFPPSOMLRTGNFTFSYTFEEVPFHSSYAHSSQSLDRLMNPLIDQ	239
Db	384	NNGSOAVGRSSPYCLEYFPPSOMLRTGNFTFTYTFEDVPFHSSYAHSSQSLDRLMNPLIDQ	443
QY	240	YLYLNRTONQSGSAQNKDLLFSRGSAGMSVQPKQWLPGPCYRQQRVSKTKTDNNNSNF	299
Db	444	YLYLSRTQTGGTANTQTLGFSQGGPNTMANQAKWLPGPCYRQQRVSTTTGQNNNSNF	503
QY	300	TWTGASKYNLNGRESIINPGTAMASHKDDKEDKFFPMSCGVMIFGKESAGASNTALDVMIT	359
Db	504	AWTAGTKYHLNGRNSLANPGIAMATHKDDDEERFFPSNGILIFGKQNAARDNADYSDVMLT	563
QY	360	DEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVLQGPIMAK	419
Db	564	SEEEIKTNPVATEEYGIADVNLQQNTAPQIGTVNSQGALPGMWQNRDVLQGPIMAK	623
QY	420	IPHTDGHFHPSPMLGCGFGLKNPPQILIKNTVPVANPPAEFSATKFASFITQYSTGVSV	479

Db 624 IPHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPADPPTTFNQSKLSFTQYSTGQSV 683

Qy 480 BIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLT 534

Db 684 BIEWELQKENSKRWNPEIQYTSNYKSTSVDFAVNTEGVSEPRPIGTRYLT 738

RESULT 10

Q67668

ID Q67668 PRELIMINARY; PRT; 534 AA.

AC Q67668;

DT 01-NOV-1996 (TremBLrel. 01, Created)

DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)

DE VP3.

GN VP3.

OS Goose parvovirus.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OX NCBI_TaxID=38251;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Virulent B;

RX MEDLINE=96010229; PubMed=7571426;

RA Zadori Z., Stefancsik R., Rauch T., Kisary J.;

RT "Analysis of the complete nucleotide sequences of goose and muscovy

RT duck parvoviruses indicates common ancestral origin with adeno-

RT associated virus 2.";

RL Virology 212:562-573(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Virulent B;

RA Zadori Z.;

RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U25749; AAA83232.1; -.

DR InterPro; IPR001403; Parvo coat.

DR Pfam; PF00740; Parvo coat; 1.

SQ SEQUENCE 534 AA; 59996 MW; F1F049558EEACE92 CRC64;

Query Match 58.4%; Score 1698; DB 12; Length 534;

Best Local Similarity 57.1%; Pred. No. 4.5e-118;

Matches 310; Conservative 87; Mismatches 128; Indels 18; Gaps 8;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNHLYKQISS 60

Db 1 MAEGGGGAMGDSGGADGVGNASGNWHCDSQWMGNTVITKTRTWLPSYNNHIYKAITS 60

Qy 61 ASTGASNDNH--YFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQ 118

Db 61 ---GTSQDANVQYAGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKSLKFIENVQ 117

Qy 119 VKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLT 178

Db 118 VKEVTTQDQTKTIANNLTSTIQVTDDEHQLPYVLGSATEGTMPPFPSSDVYALPQYGYCT 177

Qy 179 LN---NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPFHSYAHQSQDLRLMNP 235

Db 178 MHTNQNGARFNDRSAFYCLEYFPSPQMLRTGNFTFSYTFEEVPFHSMEFHSQDLRLMNP 237

Qy 236 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQORV-SKTKTDN 294

Db 238 LVDQYLWNFNEV-DSSRNAQ-----FKKAVKGAYGTMGRNWLPGPKFLDQVRAYTGGTD 291

Qy 295 NNSNFT-WTGASKYNLNGRESINPGTAMASHKDDKDFPMSGVMIFGKE--SAGASNT 351

Db 292 NYANWNWSNGKNVNLKDRQYLLQPGPVSAITYTEGEASSLPAQNILGIAKDPYRSGSTTA 351

Qy 352 ALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAMGALPGMWQDRDVI 411

Db 352 GISDIMVTEEQEVAPTNGVGWKPGRVTVTNEQNTTTAPTSSDLDVLGALPGMWQNRDIY 411

Qy 412 LOGPIWAKIPHTDGHFSPMLGGFGLKNPPPOILIKNTVPANPPAEFSATKFASFITQ 471

Db 412 LQGPICAKIPKTDGKFHSPNLGGFGLHNPFPQVFIKNTVPADPPVEYVHQWNSYITQ 471

Qy 472 YSTGQVSVBIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLT 531

Db 472 YSTGQCTVEMVWELRKENSKRWNPEIQFTSNFSNRTSIMFAPNETGGYVEDRLIGTRYLT 531

Qy 532 RPL 534

Db 532 QNL 534

RESULT 11

Q67667

ID Q67667 PRELIMINARY; PRT; 587 AA.

AC Q67667;

DT 01-NOV-1996 (TremBLrel. 01, Created)

DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)

DE VP2.

GN VP2.

OS Goose parvovirus.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

OX NCBI_TaxID=38251;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Virulent B;

RX MEDLINE=96010229; PubMed=7571426;

RA Zadori Z., Stefancsik R., Rauch T., Kisary J.;

RT "Analysis of the complete nucleotide sequences of goose and muscovy

RT duck parvoviruses indicates common ancestral origin with adeno-

RT associated virus 2.";

RL Virology 212:562-573(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Virulent B;

RA Zadori Z.;

RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U25749; AAA83231.1; -.

DR InterPro; IPR001403; Parvo coat.

DR Pfam; PF00740; Parvo coat; 1.

SQ SEQUENCE 587 AA; 65246 MW; 0278ESD5FE7F0423 CRC64;

Query Match 58.4%; Score 1698; DB 12; Length 587;

Best Local Similarity 57.1%; Pred. No. 5.2e-118;

Matches 310; Conservative 87; Mismatches 128; Indels 18; Gaps 8;

Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNHLYKQISS 60

Db 54 MAEGGGGAMGDSGGADGVGNASGNWHCDSQWMGNTVITKTRTWLPSYNNHIYKAITS 113

Qy 61 ASTGASNDNH--YFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQ 118

Db 114 ---GTSQDANVQYAGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKSLKFIENVQ 170

Qy 119 VKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLT 178

Db 171 VKEVTTQDQTKTIANNLTSTIQVTDDEHQLPYVLGSATEGTMPPFPSSDVYALPQYGYCT 230

Qy 179 LN---NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPFHSYAHQSQDLRLMNP 235

Db 231 MHTNQNGARFNDRSAFYCLEYFPSPQMLRTGNFTFSYTFEEVPFHSMEFHSQDLRLMNP 290

Qy 236 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQORV-SKTKTDN 294

Db 291 LVDQYLWNFNEV-DSSRNAQ-----FKKAVKGAYGTMGRNWLPGPKFLDQVRAYTGGTD 344

Qy 295 NNSNFT-WTGASKYNLNGRESINPGTAMASHKDDKDFPMSGVMIFGKE--SAGASNT 351

Db 345 NYANWNWSNGKNVNLKDRQYLLQPGPVSAITYTEGEASSLPAQNILGIAKDPYRSGSTTA 404

Qy 352 ALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAMGALPGMWQDRDVI 411

Db 405 GISDIMVTEEQEVAPTNGVGWKPGRVTVTNEQNTTTAPTSSDLDVLGALPGMWQNRDIY 464


```
QY 412 LOGPIWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQ 471
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 465 LOGPIGAKIPKTDGKFHPSPLNGLGFLHNPFPQVFIKNTVPADPPVEYVHQWNSYITQ 524
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 472 YSTGOVSVEIEWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLT 531
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 525 YSTGOCTVEMVWELRKENSKRWNPEIQFTSNFSNRTSIMPAPNETGGYVEDRLIGTRYLT 584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 RPL 534
:|
Db 585 QNL 587

RESULT 12
Q67666 PRELIMINARY; PRT; 732 AA.
AC Q67666;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VP1.
GN VP1.
OS Goose parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=38251;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Virulent B;
RX MEDLINE=96010229; PubMed=7571426;
RA Zadori Z., Stefancsik R., Rauch T., Kisary J.;
RT "Analysis of the complete nucleotide sequences of goose and muscovy
RT duck parvoviruses indicates common ancestral origin with adeno-
RT associated virus 2.";
RL Virology 212:562-573(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Virulent B;
RA Zadori Z.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U25749; AAA83230.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 732 AA; 81340 MW; 339507C61D47B52C CRC64;
```

```
Query Match 58.4%; Score 1698; DB 12; Length 732;
Best Local Similarity 57.1%; Pred. No. 7.1e-118;
Matches 310; Conservative 87; Mismatches 128; Indels 18; Gaps 8;

QY 1 MASGGAPMADNNEGADVGNASGNWHCDSTWLGLDRVITTTSTRTWALPTYNNHLYKQISS 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 MAEGGGAMGDSGGADVGNASGNWHCDSQWMGNTVITKTRTWLPSYNNHIYKAITS 258
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 ASTGASNDNH--YFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQ 118
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 ---GTSQDANVQYAGYSTPWGYFDNRFCHFSPRDWQRLINNNHWGIRPKSLKFKIENVQ 315
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 119 VKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLT 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 VKEVTTQDQTKTIANNLTSTIQVFTDDEHQLPYVLGSATEGMTMPFPSPDVYALPOYGYCT 375
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 179 LN---NGSQAVGRSSFFCYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 235
|||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 MHTNQNGARENDRSAFYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 435
|||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 236 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRV-SKTKTDN 294
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 436 LVDQYLWNFNEV-DSSRNAQ-----FKKAVKGAYGTMGRNWLPGPKFLDQVRAYTGGTD 489
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 295 NNSNFT-WTGASKYNLNGRESINPGTAMASHKODEDKFPMSGVMIFGKE--SAGASNT 351
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 NYANWNIWSNGKNVNLKDRQYLLQPGVPVSATYTEGEASSLPQAQNILGIADPYRSGSTTA 549
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 352 ALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPTATGDVHAMGALPGMVWQDRDVY 411
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 550 GISDIMVTEEQEVAPTNGVGWPKPYGRVTVTNEQNTTAPTSSDDLVLGALPGMVWQNRDIY 609
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 412 LOGPIWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQ 471
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 LOGPIGAKIPKTDGKFHPSPLNGLGFLHNPFPQVFIKNTVPADPPVEYVHQWNSYITQ 669
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 472 YSTGOVSVEIEWELOKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLT 531
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 YSTGOCTVEMVWELRKENSKRWNPEIQFTSNFSNRTSIMPAPNETGGYVEDRLIGTRYLT 729
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 RPL 534
:|
Db 730 QNL 732

RESULT 13
Q8V395 PRELIMINARY; PRT; 732 AA.
AC Q8V395;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Capsid protein VP.
GN VP.
OS Goose parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=38251;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPV-YG;
RA Ge Y., You Y., Xu Q.;
RT "Analysis of the major open reading frames' nucleotide sequences in
RT Goose parvovirus GPV-YG strain isolated in China.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416726; AAL37722.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 732 AA; 81456 MW; 73F2E4BC769744B6 CRC64;

Query Match 58.4%; Score 1697; DB 12; Length 732;
Best Local Similarity 57.3%; Pred. No. 8.4e-118;
Matches 311; Conservative 82; Mismatches 132; Indels 18; Gaps 8;

QY 1 MASGGAPMADNNEGADVGNASGNWHCDSTWLGLDRVITTTSTRTWALPTYNNHLYKQISS 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 MAEGGGALGDASGADVGNASGNWHCDSQWMGNTVITKTRTWLPSYNNHIYKAITS 258
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 ASTGASNDN--HYFGYSTPWGYFDNRFCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQ 118
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 ---GTSQDATVQYAGYSTPWGYFDNRFCHFSPRDWQRLINNNHWGIRPKSLKFKIENVQ 315
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 119 VKEVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLT 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 VKEVTTQDQTKTIANNLTSTIQVFTDDEHQLPYVLGSATEGMTMPFPSPDVYALPOYGYCT 375
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 179 LN---NGSQAVGRSSFFCYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 235
|||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 MHTNQNGARENDRSAFYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHQSQSLDRLMNP 435
|||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 236 LIDQYLYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRV-SKTKTDN 294
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 436 LVDQYLWNFNEV-DSSRKAQ-----FKKAVKGAYGTMGRNWLPGPKLLDQVRAYTGGTD 489
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 295 NNSNFT-WTGASKYNLNGRESINPGTAMASHKODEDKFPMSGVMIFGKE--SAGASNT 351
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 NYANWNIWSNGKNVNLKDRQYLLQPGVPVSATHTKVEASSIPAQNILGLAKDPYRSGSTTA 549
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 352 ALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPTATGDVHAMGALPGMVWQDRDVY 411
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 GISDIMVTEEQEVAPTNGVGWPKPYGRVTVTNEQNTTAPTSSDDLVLGALPGMVWQNRDIY 609
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 412 LOGPIWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFASFITQ 471
||:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
||||| 610 LQGPWAKIPKTDGKHPSPNLGGFLYNPPQVFIKNTVPADPPVEYVHOKWNSYITQ 669
QY 472 YSTQVSVIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLT 531
Db 670 YSSGQCTVMVWELKNSKRWNPFIQTSNFSNDRSIMPAPNETGGYIEDRLIGTRYLT 729
QY 532 RPL 534
Db 730 QNL 732

RESULT 14
O41855 PRELIMINARY; PRT; 734 AA.
AC O41855;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Capsid.
OS Adeno-associated virus 4.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=57579;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC VR-646;
RX MEDLINE=97404695; PubMed=9261407;
RA Chiorini J.A., Yang L., Liu Y., Safer B., Kotin R.M.;
RT "Cloning of adeno-associated virus type 4 (AAV4) and generation of
RT recombinant AAV4 particles.";
RL J. Virol. 71:6823-6833(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC VR-646;
RA Chiorini J.A., Yang L., Kotin R.M., Safer B.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U89790; AAC58045.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 734 AA; 80639 MW; 616CC27A777BBE6F CRC64;
```

```
Query Match 58.2%; Score 1690.5; DB 12; Length 734;
Best Local Similarity 59.4%; Pred. No. 2.6e-117;
Matches 325; Conservative 65; Mismatches 132; Indels 25; Gaps 9;

QY 2 ASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRWALPTYNHLYKQISSA 61
Db 199 AAAGGAIV-EGGQAGDGVGNASGDWHCDSTWSEGHVTTTSTRWVLPNTYNHLYKRLGE- 256
QY 62 STGASNDNHFGYSTPWGYFDNRFCHFSPRDWORLNNWGFPRKRLNFKLFNIQVKE 121
Db 257 ---SLQSNTYNGFSTPWGYFDNRFCHFSPRDWORLNNWGNRPKAMRVKIFNIQVKE 313
QY 122 VTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGY---LT 178
Db 314 VTTSNGETTVANNLTSTVQIFADSSYELPYVMDAQEGSLPPFPNDVFMVPQYGYCGLVT 373
QY 179 LNNGSAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNPID 238
Db 374 GNTSQOQTRDNAFCYCLEYFPPSQMLRTGNNFEITYSFEKVPFHSYAHSQSLDRLMNPID 433
QY 239 QYLYLNRNTQN---QSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN 294
Db 434 QYLMGLQSTTGTTLNAGTATTN---FTKLRPNTFSNFKKNWLPGPSIKQQGFSKTA--N 488
QY 295 NNSNFTWTGAS---KY----NLNGRESIINPGTAMASHKDDDEKFFPMSGVMIFGKESAG 347
Db 489 QNYKIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSK-FSNSQLIFAGPKQNG 547
QY 348 ASNTALDNVMTDDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQD 407
Db 548 NTATVPGLTIFTSEEELAAATNATDMDWGNLPGGDQSNLSLPTVDRLTALGAVPGMWQON 607
```

```
QY 408 RDVYLOQPIWAKIPHTDGHFHPSPMLGGFGLKNPPPOILLIKNTVPANPPAEFSATKFPAS 467
Db 608 RDIYYQGPWAKIPHTDGHFHPSPPLIGGFGLKHPPPOIFIKNTVPANPATTFSSTPVS 667
QY 468 FITQYSTQVSVIEWELQKNSKRWNPVEQYTSNYAKSANVDFTVDNNGLYTEPRPIGT 527
Db 668 FITQYSTQVSVQIDWEIQERSKRWNPVEQFTSNYQQNSLLWAPDAAGKYTEPRAIGT 727

QY 528 RYLTRPL 534
Db 728 RYLTHHL 734

RESULT 15
Q83290 PRELIMINARY; PRT; 732 AA.
AC Q83290;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Capsid protein.
OS Muscovy duck parvovirus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=37325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=89384/France;
RX MEDLINE=96406928; PubMed=8811015;
RA Le Gall-Recule G., Jestin V., Chagnaud P., Blanchard P., Jestin A.;
RT "Expression of muscovy duck parvovirus capsid proteins (VP2 and VP3)
RT in a baculovirus expression system and demonstration of immunity
RT induced by the recombinant protein.";
RL J. Gen. Virol. 77:2159-2163(1996).
DR EMBL; Z68272; CAA92575.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT CHAIN 146 732 VP2 CAPSID PROTEIN.
FT CHAIN 199 732 VP3 CAPSID PROTEIN.
SQ SEQUENCE 732 AA; 81364 MW; DE70DCCAB215F4E2 CRC64;
```

```
Query Match 58.2%; Score 1690; DB 12; Length 732;
Best Local Similarity 57.3%; Pred. No. 2.8e-117;
Matches 310; Conservative 76; Mismatches 141; Indels 14; Gaps 6;

QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRWALPTYNHLYKQISS 60
Db 199 MAEGGSGAMGDSAGGADGVGNASGNWHCDSQWLGDVITKTRTWLPSYNNHIYKAITS 258
QY 61 ASTGASNDNHFGYSTPWGYFDNRFCHFSPRDWORLNNWGFPRKRLNFKLFNIQVK 120
Db 259 GTNPDSN-TQYAGYSTPWGYFDNRFCHFSPRDWORLNNHWGIRPKALKFKIFNVQVK 317
QY 121 EVTTNDGVTTIANNLSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTN 180
Db 318 EVTTQDQTKTIANNLTSTIQIFTDNEHQLPYVLGSALEGTMPPFPSPDYALPQYGYCTMH 377
QY 181 ---NGSAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSQSLDRLMNPIL 237
Db 378 TNQSGARFNDRSAFCYCLEYFPPSQMLRTGNNFEFSFEFEVPHSMFAHSQDLDRLMNPIL 437
QY 238 QYLYLNRNTQNQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTK--TDNN 295
Db 438 QYLYWNFSEV-NGGRNAQ-----FKKAVKAGFAGMGRNWLPGPKLLDQVRAYSGGTDNY 491
QY 296 NSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMSGVMIFGKE--SAGASNTAL 353
Db 492 ANWSIWSKGNKVKFLKDREYLLQPGPVATTHTEDQASSVPAQNIIGIAKDPYRSGSTLAGI 551
QY 354 DNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDVYLO 413
Db 552 SDIMVTDEQEIAPTNGVGRWPYGLTVTNEQNTTAPTNAEVLGALPGMWQNRDIYLO 611
QY 414 GPIWAKIPHTDGHFHPSPMLGGFGLKNPPPOILLIKNTVPANPPAEFSATKFPASFITQYS 473
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:01:27 ; Search time 12.2857 Seconds
(without alignments)
1839.046 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2481.5	85.4	735	4 US-09-321-589-1	Sequence 1, Appli
2	1690.5	58.2	544	4 US-09-532-594B-18	Sequence 18, Appl
3	1690.5	58.2	598	4 US-09-532-594B-16	Sequence 16, Appl
4	1690.5	58.2	734	4 US-09-532-594B-4	Sequence 4, Appli
5	491.5	16.9	756	4 US-09-438-268-4	Sequence 4, Appli
6	479.5	16.5	543	3 US-08-856-841-22	Sequence 22, Appl
7	326	11.2	415	3 US-08-856-841-20	Sequence 20, Appl
8	318	10.9	395	3 US-08-856-841-13	Sequence 13, Appl
9	318	10.9	486	3 US-08-856-841-19	Sequence 19, Appl
10	317.5	10.9	500	3 US-08-856-841-16	Sequence 16, Appl
11	317.5	10.9	501	3 US-08-856-841-18	Sequence 18, Appl
12	304.5	10.5	264	3 US-08-856-841-14	Sequence 14, Appl
13	303	10.4	398	3 US-08-856-841-21	Sequence 21, Appl
14	209.5	7.2	579	6 5223424-13	Patent No. 5223424
15	183	6.3	584	3 US-09-022-949-2	Sequence 2, Appli
16	171	5.9	387	3 US-08-856-841-17	Sequence 17, Appl
17	119	4.1	1095	4 US-09-107-532A-3855	Sequence 3855, Ap
18	111	3.8	655	1 US-08-469-202-27	Sequence 27, Appl
19	111	3.8	655	2 US-08-484-434C-34	Sequence 34, Appl
20	111	3.8	655	4 US-09-384-361-34	Sequence 34, Appl
21	109	3.8	1186	1 US-08-485-568A-4	Sequence 4, Appli
22	109	3.8	1186	1 US-08-357-698-6	Sequence 6, Appli
23	109	3.8	1186	2 US-08-590-554A-4	Sequence 4, Appli
24	109	3.8	1186	2 US-09-184-223-4	Sequence 4, Appli
25	109	3.8	1186	5 PCT-US93-12682-6	Sequence 6, Appli
26	108	3.7	624	3 US-08-947-965-78	Sequence 78, Appl
27	108	3.7	655	1 US-08-469-202-28	Sequence 28, Appl

28	108	3.7	655	2 US-08-484-434C-35	Sequence 35, Appl
29	108	3.7	655	4 US-09-384-361-35	Sequence 35, Appl
30	104.5	3.6	717	4 US-09-626-589-1	Sequence 1, Appli
31	104	3.6	1013	3 US-09-415-522-8	Sequence 8, Appli
32	102	3.5	10182	4 US-09-134-001C-3159	Sequence 3159, Ap
33	99.5	3.4	2736	4 US-09-252-991A-30227	Sequence 30227, A
34	99	3.4	824	4 US-09-626-589-3	Sequence 3, Appli
35	98.5	3.4	997	1 US-08-232-540-1	Sequence 1, Appli
36	98.5	3.4	997	1 US-08-428-949A-1	Sequence 1, Appli
37	98.5	3.4	997	1 US-08-428-948A-1	Sequence 1, Appli
38	98.5	3.4	997	2 US-08-428-946-1	Sequence 1, Appli
39	98.5	3.4	997	5 PCT-US95-04656-1	Sequence 1, Appli
40	98.5	3.4	998	1 US-08-233-008A-6	Sequence 6, Appli
41	98.5	3.4	1000	4 US-09-193-562D-30	Sequence 30, Appl
42	98.5	3.4	1021	1 US-08-233-008A-2	Sequence 2, Appli
43	98	3.4	856	4 US-09-328-352-7255	Sequence 7255, Ap
44	98	3.4	1178	6 5254799-5	Patent No. 5254799
45	97	3.3	540	3 US-08-687-580B-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.4%; Score 2481.5; DB 4; Length 735;
Best Local Similarity 83.3%; Pred. No. 1.6e-218;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY	1	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTVNNHLYKQISS	60
Db	203	MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTVNNHLYKQISS	262
QY	61	ASTGASNDNHFGYSTPWGYFDENRHFCHFSPRDQWLNNNWGFRPKLNFNIQVK	120
Db	263	QS-GASNDNHFGYSTPWGYFDENRHFCHFSPRDQWLNNNWGFRPKLNFNIQVK	321
QY	121	EVTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPOYGYLTN	180
Db	322	EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPOYGYLTN	381
QY	181	NGSAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEEVPHSSYAHSQSLDRLMNPIDQY	240
Db	382	NGSAVGRSSFYCLEYFPPSQMLRTGNFTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY	441
QY	241	LYLNRNTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT	300
Db	442	LYLSRTNTPSGTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS	501
QY	301	WTGASKYNLNGRESIINPGTAMASHKDDKDFPPMSGVMI FGKESAGASNTALDNVMTD	360
Db	502	WTGATKYHLNGRDSLVPNPGPAMASHKDDKDFPPQSGVLIFGKQSGSEKTVVDIEKVMITD	561
QY	361	EEEEKATNPVATERFGTVAVNFQSSSTDPATGCDVHAMGALPGMVWQDRDVLQGPWAKI	420
Db	562	EEERTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI	621

QY 421 PHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLGGFGLKHPPPPQILIKNTVPANPPSTTFSAAKFASFITQYSTGQVSVE 681
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENSKRWNPEIQYTSNYKSNVNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 2
US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18

Query Match 58.2%; Score 1690.5; DB 4; Length 544;
Best Local Similarity 59.4%; Pred. No. 2.8e-146;
Matches 325; Conservative 65; Mismatches 132; Indels 25; Gaps 9;
QY 2 ASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISSA 61
Db 9 AAAGGAHV-EGGQAGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLTPTYNHLYKRLGE- 66
QY 62 STGASNDNHFGYSTPWGYFDNRFHCHFSRPRDQRLNINNNGWFRPKRLNFKLFIQVKE 121
Db 67 ---SLQSNTRYNGFSTPWGYFDNRFHCHFSRPRDQRLNINNNGWFRPKAMRVKIFNIQVKE 123
QY 122 VTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFEMIPOYGY---LT 178
Db 124 VTTSGEFTVANNLTSTVQIFADSSYELPYVMDAQEGSLPPFPNDVFMVPOYGYCGLVT 183
QY 179 LNNGSAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSSQSLDRLMNP 238
Db 184 GNTSQQTDRNATFYCYCLEYFPPSQMLRTGNNFEITYSFEKVPFHSYAHSSQSLDRLMNP 243
QY 239 QYLYYLNRTQN---QSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN 294
Db 244 QYLVGLQSTTTGTTLNAGTATTN---FTKLRPTNFSNFKNWLPGPSIKQQGFSKTA--N 298
QY 295 NNSNFTWTGAS---KY---NLNGRESINPGTAMASHKDDKDFPFMSGVMI FGKESAG 347
Db 299 QNYKIPATGSDSLIKYETHSTLDGRWSALTTPGPPMATAGPADSK-FSNSQLIFAGPKQNG 357
QY 348 ASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQD 407
Db 358 NTATVPGLTIFTSEELAAATNATDMDWGNLPGGDQSNLSNLTVDRLTALGAVPGMVWQN 417
QY 408 RDVYLGQPIWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFA 467
Db 418 RDIYYQGPIWAKIPHTDGHFHPSPMLGGFGLKHPPPQIFIKNTVPANPATTFSSTPVNS 477
QY 468 FITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGT 527

Db 478 FITQYSTGQVSQVIDWEIQERSKRWNPEVQFTSNYQQNSLLWAPDAAGKYTEPRAIGT 537
QY 528 RYLTRPL 534
Db 538 RYLTHHL 544
RESULT 3
US-09-532-594B-16
; Sequence 16, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16

Query Match 58.2%; Score 1690.5; DB 4; Length 598;
Best Local Similarity 59.4%; Pred. No. 3.2e-146;
Matches 325; Conservative 65; Mismatches 132; Indels 25; Gaps 9;
QY 2 ASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISSA 61
Db 63 AAAGGAHV-EGGQAGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLTPTYNHLYKRLGE- 120
QY 62 STGASNDNHFGYSTPWGYFDNRFHCHFSRPRDQRLNINNNGWFRPKRLNFKLFIQVKE 121
Db 121 ---SLQSNTRYNGFSTPWGYFDNRFHCHFSRPRDQRLNINNNGWFRPKAMRVKIFNIQVKE 177
QY 122 VTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGLPPFPADVFEMIPOYGY---LT 178
Db 178 VTTSGEFTVANNLTSTVQIFADSSYELPYVMDAQEGSLPPFPNDVFMVPOYGYCGLVT 237
QY 179 LNNGSAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHSSQSLDRLMNP 238
Db 238 GNTSQQTDRNATFYCYCLEYFPPSQMLRTGNNFEITYSFEKVPFHSYAHSSQSLDRLMNP 297
QY 239 QYLYYLNRTQN---QSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN 294
Db 298 QYLVGLQSTTTGTTLNAGTATTN---FTKLRPTNFSNFKNWLPGPSIKQQGFSKTA--N 352
QY 295 NNSNFTWTGAS---KY---NLNGRESINPGTAMASHKDDKDFPFMSGVMI FGKESAG 347
Db 353 QNYKIPATGSDSLIKYETHSTLDGRWSALTTPGPPMATAGPADSK-FSNSQLIFAGPKQNG 411
QY 348 ASNTALDNVMTDEEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQD 407
Db 412 NTATVPGLTIFTSEELAAATNATDMDWGNLPGGDQSNLSNLTVDRLTALGAVPGMVWQN 471
QY 408 RDVYLGQPIWAKIPHTDGHFHPSPMLGGFGLKNPPPPQILIKNTVPANPPAEFSATKFA 467
Db 472 RDIYYQGPIWAKIPHTDGHFHPSPMLGGFGLKHPPPQIFIKNTVPANPATTFSSTPVNS 531
QY 468 FITQYSTGQVSVEIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGT 527
Db 532 FITQYSTGQVSQVIDWEIQERSKRWNPEVQFTSNYQQNSLLWAPDAAGKYTEPRAIGT 591

IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
PUBLICATION RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 1982
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
PUBLICATION RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
PUBLICATION RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13

Query Match 10.9%; Score 318; DB 3; Length 395;
Best Local Similarity 25.4%; Pred. No. 8.8e-21;
Matches 99; Conservative 61; Mismatches 167; Indels 62; Gaps 13;
QY 128 VTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTNN-GSQAV 186
DB 1 MTMITPSLHACMLV--DHEYKYPYVLGGQDTLAPELPIWVYFPQYAYLTGVDVNTQGI 58
QY 187 G-----RSSFYCLEYFPFQMLRTGNNFTFSYTFEEVPHSSYAHSSQSLDRLMNP 237
DB 59 SGDSKKLASEESAFVLEHSSFQLLGTGCTASMSYKFPVPVPPENLEGCSQHFMVYNPL- 117
QY 238 DOYLYLNRNTQNSGSAQNKDLLFSRGSPPAGMSVQPKWLPGPCYRQQRVSKTKTDNNNS 297
DB 118 --YGSRLGVDPDTLGGDPKFRSL-----THEDHAIQPNFMFGPLVNSVSTKEGDSNTGA 170
QY 298 NFTWTGASKYNLNGRESIINPG-TAMASHKODEDKFPFMSGVMIFGKESAGASNTALDNV 356
DB 171 GKALTGLSTGTSQNTSRISLRPGVSPQPYHHWDTDKYVTGAINAISHGQTTYG---NAEDKE 227
QY 357 -----MITDEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRD 409
DB 228 YQGVGRFPNEKEQLKQLQGLNMHTY-----FPNKGQTYTDQIE-RPLMVGSVWNRRA 280
QY 410 VYLQGPWAKIPHTDGHFHP-PLMGGFGLKNPPQILIKNTVPANPPAEFSATKFASE 468

DB 281 LHYESQLWSKIPNLDSDSKTQFAALGGWGLHQPPQI-----P 318
QY 469 ITQYSTGOVSVEIEWEL-QKENSKRWNPE 496
DB 319 LKQYAVGIMTVMFTFKLGPBKATGRWNPO 347
RESULT 9
US-08-856-841-19
Sequence 19, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 486
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.

CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 264
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 14:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:

PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 14:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 14:
US-08-856-841-14
Query Match 10.5%; Score 304.5; DB 3; Length 264;
Best Local Similarity 32.5%; Pred. No. 8.2e-20;
Matches 76; Conservative 30; Mismatches 105; Indels 23; Gaps 5;
Qy 7 APVADNNEGADVGNASGNWCHDSTWLGDRVITTTSTRTWALPTNNHLYKQISSASTGAS 66
Db 9 AAEASTGAGGGGNSVKSMWSEGFATFSAVCTFTRQFLIPYDPEHHYKVFSPAASSCH 68
Qy 67 ND-----NHYFGYSTPWGYDFNRFHCHFSRDRWRLNNWGFPRKRLNFKLEN 116
Db 69 NASGKEAKVCTISPINGYSTPWRYLDFNALNLFSPLEFQHLIENYGSIAPDALTVTISE 128
Qy 117 IQVKEVT--TNDGVTTIANLNTSTVQVFSDSYQLPYVLGSAHQGCLPPFPADVEMIPQY 174
Db 129 IAVKDVTDKTGGV-QVTDSTTGRCLMLVDHEYKYPYVLGGQDTLAPELPIWYFPPQY 187
Qy 175 GYLTLNN-GSQAVG-----RSSFYCLEYFPSPQMLRTGNNFTFSYTFEEVP 218
Db 188 AYLTVGDVNTQGISGDSKKLASESAFYVLEHSSFQLLGTGTASMSYKFPVP 241
RESULT 13
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994


```

Db      393  RYTFNPQSKGRAPKQFNNQAPLNLNTNNGTLLPSDPIGGKSNKHFMTLNTYGPLTA 455
QY      404  VWQDRDVLQGPWAKIPHTD--GHFHPSPIMGGFGLK-NPPPQILIKNTVPVPANPPAEF 460
Db      453  LNNTPAVFPNGQIWDKELDTDLKPRLH---VTAPFVCKNNPPGQLFVKIAP---NLTDDE 506
QY      461  SA--TKFASFITQYSTQGVSVIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGL 518
Db      507  NADSPQPPRIITD-SNFWWKGLTFTAKWRSSNMWNPIQQHTT-----TAENIRK 555
QY      519  YTEPRPIG 526
Db      556  YI-PTNIG 562

RESULT 15
US-09-022-949-2
; Sequence 2, Application US/09022949
; Patent No. 6187759
; GENERAL INFORMATION:
; APPLICANT: Tarpey, Ian
; APPLICANT: Greenwood, Neil
; TITLE OF INVENTION: Canine parvovirus DNA vaccination
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 6187759el Patent Dept.
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30(EPC)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,949
; FILING DATE: 11-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-022-949-2

```

Query Match	6.3%;	Score 183;	DB 3;	Length 584;
Best Local Similarity	20.4%;	Pred. NO. 3.7e-08;		
Matches 130;	Conservative 91;	Mismatches 223;	Indels 192;	Gaps 33;

Qy	5	GGAPMADNNEGADGVGNASGNWHCDSTLWLGDRVITTTST-----RTWALPTYNN-	52
		: : : : :	
Db	10	GGQP-AVRNERATGSGNGSGGGGGGS--GGVGISTGTENNCTEFKFLENGWVEITANSS	66
		: : : : :	
Qy	53	---HL-----YKQI-----SSASTGASNDNHYPGYSTPWGYFDENRFCHCHSPRDW	95
		: : : : : : :	
Db	67	RLVHLNMPESENYRRVVVNNLDKTAVNGNMLDTHAQIVTPWSLVDANAWGVWFNPGDW	126
		: : : : :	
Qy	96	QRLINNNGFRPKRLNFKLFNIQVKEVT--TNDGVTTIANNLSTVQVFSDSQVQLPYV	152
		: : : : : : :	
Db	127	QLIVNTMSSELHVSFEQEIFNVVLKTVSESATQPPTKVYNNDLTASLMVALDSNNIMPFT	186
		: : : : :	
Qy	153	LGSAHQGCL-----PPFPA-----DVFMIPQYGYLTNNGSQAVGRS-----	189
		: : : : : : :	

DB	187	PAAMRSETLGFYPWKPTIPTPWRYFFQWDRTLIPSH-----TGTSGTPTNIYHGTDPDD	240
QY	190	-SFYCLE-YFPSQMLRTGNPF-TFSYTFEEVFPFHSSYAHQSOLDRLMNPPLIDQYLYYNR	246
DB	241	VQFYTIENSVPVHLRLTGDEFATGTFFFDCKP--CRLTHWTQNRALG--LPPFLNSLPQ	296
QY	247	TQNQS-----GSAQNK-----DLLFSRGSPPAGMSVQPKNWLPGPCYRQQRV	287
DB	297	SEGGTNEFYIGVQQDKRGVTQMGTNTNYITEATIMRPAEUVGYS-----APYYSFEAS	348
QY	288	SK-----TKTDNNNSFTWTGASKY---NLNGRESIINPGT---AMASHKD	327
DB	349	TQGPFKTPIAAGRGGAQTDENQAA---DGDPRYAFGRHQGKTTTGTGETPERFTYIAHQD	405
QY	328	DEDKFFPMGVMIFGKESAGASNTALD-NVMITDEEEIKATNPVATERFGTVAVNFQSSS	386
DB	406	T-----GRYPEGDWIQNINENLPVTNDNVLLPTDPDG---GKTGINY--TN	446
QY	387	TDPATGDVHAMGALPGMWVQDRDYYLQGPIWAKIPHDTGDFHPSPLMGFGFLK-----	439
DB	447	IFNTYGPLTALNNVP-----PVYPNGQIWDKEFTD-----LKPRLHVNA	486
QY	440	-----NPPPQILLIKNTVPVAN---PPAEFSATKPFASFITQYSTQGVSVIEWELQKENS	490
DB	487	PFVCQNNCPGQLFVKVAPNLNTNEYDPPDASANMSR----IVTYSDFWWKGLVFKAKLRAS	542
QY	491	KRWNPVQYTSNYAKSANVDFTVDNNGLYTEPRPIG	526
DB	543	HTWNPIQQMSIN-----IDNQFNYYV-PSNIG	567

Search completed: January 21, 2004, 16:08:03
Job time : 13.2857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 21, 2004, 16:04:27 ; Search time 24 Seconds
(without alignments)
4549.706 Million cell updates/sec

Title: US-09-807-802A-17
Perfect score: 2906
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPICTRYLTRPL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	736	12	US-10-291-583-64
2	2906	100.0	736	12	US-10-423-704A-5
3	2884	99.2	736	12	US-10-291-583-65
4	2528	87.0	736	12	US-10-291-583-71
5	2528	87.0	736	12	US-10-423-704A-6
6	2481.5	85.4	533	14	US-10-038-972A-15
7	2481.5	85.4	598	14	US-10-038-972A-14
8	2481.5	85.4	735	12	US-10-291-583-70
9	2481.5	85.4	735	12	US-10-423-704A-4
10	2481.5	85.4	735	14	US-10-038-972A-13
11	2481.5	85.4	735	15	US-10-293-478-1
12	2451.5	84.4	735	12	US-10-291-583-68
13	2449.5	84.3	735	12	US-10-291-583-67
14	2448.5	84.3	735	12	US-10-291-583-66
15	2445.5	84.2	735	12	US-10-291-583-69

16	2441	84.0	728	12	US-10-291-583-106	Sequence 106, Appl
17	2439	83.9	728	12	US-10-291-583-102	Sequence 102, Appl
18	2437.5	83.9	731	12	US-10-291-583-89	Sequence 89, Appl
19	2436	83.8	728	12	US-10-291-583-103	Sequence 103, Appl
20	2436	83.8	728	12	US-10-291-583-107	Sequence 107, Appl
21	2436	83.8	728	12	US-10-291-583-108	Sequence 108, Appl
22	2434.5	83.8	738	12	US-10-291-583-93	Sequence 93, Appl
23	2434.5	83.8	738	12	US-10-291-583-94	Sequence 94, Appl
24	2431.5	83.7	731	12	US-10-291-583-88	Sequence 88, Appl
25	2430	83.6	737	12	US-10-291-583-2	Sequence 2, Appl
26	2430	83.6	737	12	US-10-423-704A-8	Sequence 8, Appl
27	2429.5	83.6	738	12	US-10-291-583-79	Sequence 79, Appl
28	2429.5	83.6	738	12	US-10-291-583-81	Sequence 81, Appl
29	2428.5	83.6	738	12	US-10-291-583-85	Sequence 85, Appl
30	2428.5	83.6	738	12	US-10-291-583-91	Sequence 91, Appl
31	2427.5	83.5	738	12	US-10-291-583-92	Sequence 92, Appl
32	2425	83.4	728	12	US-10-291-583-105	Sequence 105, Appl
33	2423.5	83.4	733	12	US-10-291-583-87	Sequence 87, Appl
34	2422.5	83.4	733	12	US-10-291-583-86	Sequence 86, Appl
35	2422.5	83.4	738	12	US-10-291-583-80	Sequence 80, Appl
36	2422	83.3	728	12	US-10-291-583-101	Sequence 101, Appl
37	2421.5	83.3	733	12	US-10-291-583-90	Sequence 90, Appl
38	2421.5	83.3	738	12	US-10-291-583-82	Sequence 82, Appl
39	2419.5	83.3	729	12	US-10-291-583-110	Sequence 110, Appl
40	2418.5	83.2	729	12	US-10-291-583-111	Sequence 111, Appl
41	2418	83.2	728	12	US-10-291-583-104	Sequence 104, Appl
42	2417	83.2	737	12	US-10-291-583-72	Sequence 72, Appl
43	2415.5	83.1	738	12	US-10-291-583-84	Sequence 84, Appl
44	2413	83.0	736	12	US-10-291-583-100	Sequence 100, Appl
45	2413	83.0	736	12	US-10-423-704A-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-10-291-583-64
; Sequence 64, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV1
US-10-291-583-64

Query Match 100.0%; Score 2906; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 2.2e-271;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRWTALPTYNHLYKQISS 60
Db	203	MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRWTALPTYNHLYKQISS 262
Qy	61	ASTGASNDNHFGYSTPWGFDNRFCHFSRDNQRLNNNNGWFRKRLNFKLFNQVK 120

Db 263 ASTGASNDNHFGYSTPWGYDFNRFCHFSRDPWQRLNINNWGFRPKRLNFKLFNIQVK 322
Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 180
Db 323 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 382
Qy 181 NGSQAVGRSSFCLEYFSPQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPIDQY 240
Db 383 NGSQAVGRSSFCLEYFSPQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPIDQY 442
Qy 241 LYTLNRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 443 LYTLNRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNVMITD 360
Db 503 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNVMITD 562
Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVYLGQPIWAKI 420
Db 563 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVYLGQPIWAKI 622
Qy 421 PHTDGHFSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 623 PHTDGHFSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 682
Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 683 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736

RESULT 2
US-10-423-704A-5
; Sequence 5, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-423-704A-5

Query Match 100.0%; Score 2906; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 2.2e-271;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITTSRTWALPTYNNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITTSRTWALPTYNNHLYKQISS 262
Qy 61 ASTGASNDNHFGYSTPWGYDFNRFCHFSRDPWQRLNINNWGFRPKRLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPWGYDFNRFCHFSRDPWQRLNINNWGFRPKRLNFKLFNIQVK 322

Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 180
Db 323 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 382
Qy 181 NGSQAVGRSSFCLEYFSPQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPIDQY 240
Db 383 NGSQAVGRSSFCLEYFSPQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPIDQY 442
Qy 241 LYTLNRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 443 LYTLNRTQNSGSAQNKDLLFSRGSAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 502
Qy 301 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNVMITD 360
Db 503 WTGASKYNLNGRESIIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNVMITD 562
Qy 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVYLGQPIWAKI 420
Db 563 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMWQDRDVYLGQPIWAKI 622
Qy 421 PHTDGHFSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 623 PHTDGHFSPMLMGFGGLKNPPQILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 682
Qy 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 683 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 736

RESULT 3
US-10-291-583-65
; Sequence 65, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences and Isolating No. US20030138772A1
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 736
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV6VP1
US-10-291-583-65
Query Match 99.2%; Score 2884; DB 12; Length 736;
Best Local Similarity 99.1%; Pred. No. 3e-269;
Matches 529; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITTSRTWALPTYNNHLYKQISS 60
Db 203 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTVITTSRTWALPTYNNHLYKQISS 262
Qy 61 ASTGASNDNHFGYSTPWGYDFNRFCHFSRDPWQRLNINNWGFRPKRLNFKLFNIQVK 120
Db 263 ASTGASNDNHFGYSTPWGYDFNRFCHFSRDPWQRLNINNWGFRPKRLNFKLFNIQVK 322
Qy 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 180
Db 323 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGCLPPFPADVFMIPOYGYLTIN 382

Db 442 LYYLNRTQGTSGTTNQSRLLFSQAGPQSMQLQARNWLPGPCYRQORLSKTDANNNSNF 501
QY 300 TWTGASKYNLNGRESINPGTAMASHKDDKDEKFFPMGVMIFGKESAGASNTALDNVMIT 359
Db 502 PWTAAKYHLNGRDSLVPNGPAMASHKDDKDEKFFPMHGNLIFGKEGTTASNAELDNVMIT 561
QY 360 DEEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAK 419
Db 562 DEEEIRTTNPVATEQYGTVANLQSSNTAPTGTGNHOGALPGMVWQDRDVLQGPWAK 621
QY 420 IPHTDGHFHPSPMLMGFGLKNPPPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSV 479
Db 622 IPHTDGHFHPSPMLMGFGLKHPPPPQIMIKNTVPANPPPTTFSPAKEFSFITQYSTGQVSV 681
QY 480 EIEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 682 EIEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL 736

RESULT 6
US-10-038-972A-15
; Sequence 15, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 533
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP3 capsid protien
US-10-038-972A-15

Query Match 85.4%; Score 2481.5; DB 14; Length 533;
Best Local Similarity 83.3%; Pred. No. 1.4e-230;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGRVITTTSTRTWALPTYNHLYKQISS 60
Db 1 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNHLYKQISS 60
QY 61 ASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWORLINNNWGFPRKLNPKLFNIQVK 120
Db 61 QS-GASNDNHYFGYSTPWGYFDENRFCHFSPRDWORLINNNWGFPRKLNPKLFNIQVK 119
QY 121 EVTNDGVTTIANNLSTVQVFDSEYQLPYVLGSAHQGCCLPPFPADVFMIPOYGYLTLN 180
Db 120 EVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCCLPPFPADVFVMPQYGYLTLN 179
QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPHSSYAHSQSLDRLMNPIDQY 240
Db 180 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 239
QY 241 LYYLNRTQNSGSAQNKDILLFSRGSPPAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 300
Db 240 LYYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQORVSKTSADNNNSEYS 299
QY 301 WTGASKYNLNGRESINPGTAMASHKDDKDEKFFPMGVMIFGKESAGASNTALDNVMITD 360
Db 300 WTGATKYHLNGRDSLVPNGPAMASHKDDKDEKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 359
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
Db 360 EEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 419
QY 421 PHTDGHFHPSPMLMGFGLKNPPPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSV 480

Db 420 PHTDGHFHPSPMLMGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSV 479
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 480 IEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL 533
RESULT 7
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 598
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP2 capsid protien
US-10-038-972A-14

Query Match 85.4%; Score 2481.5; DB 14; Length 598;
Best Local Similarity 83.3%; Pred. No. 1.6e-230;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;
QY 1 MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGRVITTTSTRTWALPTYNHLYKQISS 60
Db 66 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNHLYKQISS 125
QY 61 ASTGASNDNHYFGYSTPWGYFDENRFCHFSPRDWORLINNNWGFPRKLNPKLFNIQVK 120
Db 126 QS-GASNDNHYFGYSTPWGYFDENRFCHFSPRDWORLINNNWGFPRKLNPKLFNIQVK 184
QY 121 EVTNDGVTTIANNLSTVQVFDSEYQLPYVLGSAHQGCCLPPFPADVFMIPOYGYLTLN 180
Db 185 EVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCCLPPFPADVFVMPQYGYLTLN 244

QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEEVPHSSYAHSQSLDRLMNPIDQY 240
Db 245 NGSQAVGRSSFYCLEYFPPSQMLRTGNNTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 304
QY 241 LYYLNRTQNSGSAQNKDILLFSRGSPPAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSNFT 300
Db 305 LYYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQORVSKTSADNNNSEYS 364
QY 301 WTGASKYNLNGRESINPGTAMASHKDDKDEKFFPMGVMIFGKESAGASNTALDNVMITD 360
Db 365 WTGATKYHLNGRDSLVPNGPAMASHKDDKDEKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 424
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
Db 425 EEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 484
QY 421 PHTDGHFHPSPMLMGFGLKNPPPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSV 480
Db 485 PHTDGHFHPSPMLMGFGLKHPPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSV 544
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 545 IEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL 598

RESULT 8
US-10-291-583-70
; Sequence 70, Application US/10291583
; Publication No. US20030138772A1

; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US2003013872A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV2
US-10-291-583-70

Query Match 85.4%; Score 2481.5; DB 12; Length 735;
Best Local Similarity 83.3%; Pred. No. 2.3e-230;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVTTSTRTWALPTYNHLYKQISS 60
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 203 MATGSGAPMADNNEGADGVGNSSGNWCHDSTWMDRVTTSTRTWALPTYNHLYKQISS 262
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 ASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263 QS-GASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 321
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTN 180
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 322 EVTQNDGTTTIANNTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTN 381
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 NGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEVPFHSSYAHQSGLDRLMPLIDQY 240
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 382 NGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEVPFHSSYAHQSGLDRLMPLIDQY 441
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 LYLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSFT 300
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 442 LYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNSEYS 501
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 WTGASKYNLNGRESINPGTAMASHKDDKEDKFFPMGSGVMIFGKESAGASNTALDNVMTD 360
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 502 WTGATKYHLNGRDSLVPNGPAMASHKDDKEDKFFPQSGVLIFGKGSEKTNVDIEKVMITD 561
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 562 EEEIRTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 621
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 PHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPPAEFSATKFSFITQYSTGQVSVE 480
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 622 PHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPPSTTFSAAKFSFITQYSTGQVSVE 681
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 682 IEWELQKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 735
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9
US-10-423-704A-4
; Sequence 4, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 2
US-10-423-704A-4

Query Match 85.4%; Score 2481.5; DB 12; Length 735;
Best Local Similarity 83.3%; Pred. No. 2.3e-230;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVTTSTRTWALPTYNHLYKQISS 60
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 203 MATGSGAPMADNNEGADGVGNSSGNWCHDSTWMDRVTTSTRTWALPTYNHLYKQISS 262
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 ASTGASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 120
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263 QS-GASNDNHFGYSTPWGYFDNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVK 321
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTN 180
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 322 EVTQNDGTTTIANNTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTN 381
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 NGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEVPFHSSYAHQSGLDRLMPLIDQY 240
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 382 NGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEVPFHSSYAHQSGLDRLMPLIDQY 441
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 LYLNRTQNSGSAQNKDLLFSRGS PAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSFT 300
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 442 LYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNSEYS 501
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 WTGASKYNLNGRESINPGTAMASHKDDKEDKFFPMGSGVMIFGKESAGASNTALDNVMTD 360
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 502 WTGATKYHLNGRDSLVPNGPAMASHKDDKEDKFFPQSGVLIFGKGSEKTNVDIEKVMITD 561
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 562 EEEIRTNPVATEQYGSVSTNLQRNQAATADVNTQGVLPGMVWQDRDVLQGPWAKI 621
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 PHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPPAEFSATKFSFITQYSTGQVSVE 480
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 622 PHTDGHFHPSPMLMGFGFLKNPPQILIKNTVPANPPSTTFSAAKFSFITQYSTGQVSVE 681
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNGLYTEPRPIGTRYLTRPL 534
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 682 IEWELQKENSKRWNPEIQYTSNYKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL 735
||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04

```

; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP1 capsid protien
US-10-038-972A-13

Query Match
Best Local Similarity 85.4%; Score 2481.5; DB 14; Length 735;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 262

QY 61 ASTGASNDNHYFGYSTPWCYFDNRFHCHFSPRDWORLNNNWGFRPKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHYFGYSTPWCYFDNRFHCHFSPRDWORLNNNWGFRPKRLNFKLFNIQVK 321

QY 121 EVTTNDGVTTIANNLTSVQVFSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 180
Db 322 EVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 381

QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 240
Db 382 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 441

QY 241 LYYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 442 LYYLSRNTPTSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 501

QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGVMIFFGKESAGASNTALDNNVMTD 360
Db 502 WTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLIFFGKQSEKTNVDIEKVMITD 561

QY 361 EEEIKATNPVATERFGTVAVNPFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWIWAKI 420
Db 562 EEEIRTTNPVATEQYGSVSTNLQRNGRQAATADVNTQGVLPGMVWQDRDVLQGPWIWAKI 621

QY 421 PHTDGHFHPSPMLMGFGGLKNPPQIILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLMGFGGLKHPPQIILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVE 681

QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 11
US-10-293-478-1
; Sequence 1, Application US/10293478
; Publication No. US20030078411A1
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match
85.4%; Score 2481.5; DB 15; Length 735;
```

```

Best Local Similarity 83.3%; Pred. No. 2.3e-230;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 262

QY 61 ASTGASNDNHYFGYSTPWCYFDNRFHCHFSPRDWORLNNNWGFRPKRLNFKLFNIQVK 120
Db 263 QS-GASNDNHYFGYSTPWCYFDNRFHCHFSPRDWORLNNNWGFRPKRLNFKLFNIQVK 321

QY 121 EVTTNDGVTTIANNLTSVQVFSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 180
Db 322 EVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTN 381

QY 181 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 240
Db 382 NGSQAVGRSSFYCLEYFPPSQMLRTGNNFTFSYTFEEVPHSSYAHQSQSLDRLMNLIDQY 441

QY 241 LYYLNRNTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 442 LYYLSRNTPTSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYS 501

QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKFFPMSGVMIFFGKESAGASNTALDNNVMTD 360
Db 502 WTGATKYHLNGRDSLVPNPGPAMASHKDDKFFPQSGVLIFFGKQSEKTNVDIEKVMITD 561

QY 361 EEEIKATNPVATERFGTVAVNPFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWIWAKI 420
Db 562 EEEIRTTNPVATEQYGSVSTNLQRNGRQAATADVNTQGVLPGMVWQDRDVLQGPWIWAKI 621

QY 421 PHTDGHFHPSPMLMGFGGLKNPPQIILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLMGFGGLKHPPQIILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVE 681

QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL 735

RESULT 12
US-10-291-583-68
; Sequence 68, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (/
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifie
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.4
US-10-291-583-68

Query Match
84.4%; Score 2451.5; DB 12; Length 735;
Best Local Similarity 82.6%; Pred. No. 1.8e-227;
Matches 441; Conservative 37; Mismatches 55; Indels 1; Gaps 1;
```

QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MASGGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPWGYFDENRFCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVK 120
Db 263 ES-GATNDNHFGYSTPWGYFDENRFCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVK 321
QY 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTIN 180
Db 322 EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTIN 381
QY 181 NGSQAVGRSSFCLEYFPPSQMLRTGNNTFSYTFEEVPHSSYAHSQSLDRLMNPIDQY 240
Db 382 NGSQAVGRSSFCLEYFPPSQMLRTGNNTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 441
QY 241 LYILNRTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 442 LYILSKTQGTSGTTQSRQLQFSQAGPSSMAQQAQKNWLPGPSYRQQRMSKTANDNNNSEFA 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNNVITD 360
Db 502 WTAATKYLLNGRNSLVNPGPPMASHKODEEKYFPMHGNLIFGKQGTGTTNVDIESTLTD 561
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 562 EEEIRTNPVATEYQGVATNHQSODTTASVGSVDSQGIPLPGMWQDRDVLQGPWAKT 621
QY 421 PHTDGHFHPSPMLGGFGLKNPPQIILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLGGFGLKHPPQIILIKNTVPANPATFTTFCGFASFITQYSTGQVSVE 681
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENSKRWNPEIQYTSNYKSNVNEFTVDANGVYSEPRPIGTRYLTRNL 735

RESULT 13
US-10-291-583-67
; Sequence 67, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.7
US-10-291-583-67

Query Match 84.3%; Score 2449.5; DB 12; Length 735;
Best Local Similarity 82.6%; Pred. No. 2.8e-227;
Matches 441; Conservative 37; Mismatches 55; Indels 1; Gaps 1;
QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MASGGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 262

QY 61 ASTGASNDNHFGYSTPWGYFDENRFCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVK 120
Db 263 ES-GATNDNHFGYSTPWGYFDENRFCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVK 321
QY 121 EVTTNDGVTTIANNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTIN 180
Db 322 EVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTIN 381
QY 181 NGSQAVGRSSFCLEYFPPSQMLRTGNNTFSYTFEEVPHSSYAHSQSLDRLMNPIDQY 240
Db 382 NGSQAVGRSSFCLEYFPPSQMLRTGNNTFSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 441
QY 241 LYILNRTQNQSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFT 300
Db 442 LYILSKTQGTSGTTQSRQLQFSQAGPSSMAQQAQKNWLPGPSYRQQRMSKTANDNNNSEFA 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKODEDKFFPMSGVMI FGKESAGASNTALDNNVITD 360
Db 502 WTAATKYLLNGRNSLVNPGPPMASHKODEEKYFPMHGNLIFGKQGTGTTNVDIESTLTD 561
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMGALPGMWQDRDVLQGPWAKI 420
Db 562 EEEIRTNPVATEYQGVATNHQSODTTASVGSVDSQGIPLPGMWQDRDVLQGPWAKT 621
QY 421 PHTDGHFHPSPMLGGFGLKNPPQIILIKNTVPANPPAEFSATKFASFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLGGFGLKHPPQIILIKNTVPANPATFTTFCGFASFITQYSTGQVSVE 681
QY 481 IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENSKRWNPEIQYTSNYKSNVNEFTVDANGVYSEPRPIGTRYLTRNL 735

RESULT 14
US-10-291-583-66
; Sequence 66, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.3
US-10-291-583-66

Query Match 84.3%; Score 2448.5; DB 12; Length 735;
Best Local Similarity 82.4%; Pred. No. 3.5e-227;
Matches 440; Conservative 38; Mismatches 55; Indels 1; Gaps 1;
QY 1 MASGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISS 60
Db 203 MASGGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISS 262
QY 61 ASTGASNDNHFGYSTPWGYFDENRFCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVK 120

Db 263 ES-GATNDNHFGYSTPWGYDFDNRFHCHFSPRDWQRLNNWGRPKLNFKLFNIQVK 321
QY 121 EVTTNDGVTTIANNLTSTVQVFDSEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTN 180
Db 322 EVTQNDGTTIANNLTSAVQVFTDSEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTN 381
QY 181 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSLSLDRMLNPLIDQY 240
Db 382 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEDVPHSSYAHQSLSLDRMLNPLIDQY 441
QY 241 LYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSFT 300
Db 442 LYLSKTQGTSGTTQOSRLQFSAQSPSSMAQQAQKWLPGPSYRQORMSKTANDNNSEFA 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKEDKFFPMSGVMI FGKESAGASNTALDNVMTD 360
Db 502 WTAATKYLYNGRNSLVNPGPPVASHKDDKEDKFFPMSGVMI FGKESAGASNTALDNVMTD 561
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
Db 562 EEEIRTNPVATEQYQGVATNRQSNQNTTASVGSVDSQGLPGMVWQDRDVLQGPWAKT 621
QY 421 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPATTTTPGKFASFITQYSTGQVSVE 681
QY 481 IEWELQKENS KRWNPVEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENS KRWNPVETVQYTSNYAKSVNVEFTVDANGVYSEPRPIGTRYLTRNL 735

RESULT 15
US-10-291-583-69
; Sequence 69, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 69
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone A3.5
US-10-291-583-69

Query Match 84.2%; Score 2445.5; DB 12; Length 735;
Best Local Similarity 82.4%; Pred. No. 6.8e-227;
Matches 440; Conservative 38; Mismatches 55; Indels 1; Gaps 1;
QY 1 MASGGGAPMADNNEGADGVGNASGNWCHDSTWLGDRTVITSTRWALPTNNHLYKOISS 60
Db 203 MASGGGAPMADNNEGADGVGNSSGNWCHDSTWMDRTVITSTRWALPTNNHLYKOISS 262
QY 61 ASTGASNDNHFGYSTPWGYDFDNRFHCHFSPRDWQRLNNWGRPKLNFKLFNIQVK 120
Db 263 ES-GATNDNHFGYSTPWGYDFDNRFHCHFSPRDWQRLNNWGRPKLNFKLFNIQVK 321
QY 121 EVTTNDGVTTIANNLTSTVQVFDSEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTN 180

Db 322 EVTQNDGTTIANNLTSTVQVFDSEYQLPYVLGSAHQCLPPFPADVFMIPQYGYLTN 381
QY 181 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEEVPHSSYAHQSLSLDRMLNPLIDQY 240
Db 382 NGSQAVGRSSFYCLEYFPSPQMLRTGNFTFSYTFEDVPHSSYAHQSLSLDRMLNPLIDQY 441
QY 241 LYLNRTQNSGSAQNKDLLFSRGSFAGMSVQPKNWLPGPCYRQORVSKTKTDNNNSFT 300
Db 442 LYLSKTQGTSGTTQOSRLQFSAQSPSSMAQQAQKWLPGPSYRQORMSKTANDNNSEFA 501
QY 301 WTGASKYNLNGRESIINPGTAMASHKDDKEDKFFPMSGVMI FGKESAGASNTALDNVMTD 360
Db 502 WTAATKYLYNGRNSLVNPGPPVASHKDDKEDKFFPMSGVMI FGKESAGASNTALDNVMTD 561
QY 361 EEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVLQGPWAKI 420
Db 562 EEEIRTNPVATEQYQGVATNRQSNQNTTASVGSVDSQGLPGMVWQDRDVLQGPWAKT 621
QY 421 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPPAEFSATKFAFITQYSTGQVSVE 480
Db 622 PHTDGHFHPSPMLGGFGLKNPPQILIKNTVPANPATTTTPGKFASFITQYSTGQVSVE 681
QY 481 IEWELQKENS KRWNPVEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL 534
Db 682 IEWELQKENS KRWNPVETVQYTSNYAKSVNVEFTVDANGVYSEPRPIGTRYLTRNL 735

Search completed: January 21, 2004, 16:18:25
Job time : 25 secs